

Mauricio Matthesius

**Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter
Daten im Information Lifecycle Management**

**Dissertation
zur Erlangung des akademischen Grades
„Doctor rerum politicarum“
an der Fakultät für Wirtschaftswissenschaften und Medien
der Technischen Universität Ilmenau**

Fachgebiet Informations- und Wissensmanagement

1. Gutachter: Prof. Dr. Dirk Stelzer

2. Gutachter: Prof. Dr. Volker Nissen

Tag der Einreichung: 12. August 2015

Tag der wissenschaftlichen Aussprache: 20. Januar 2016

URN: urn:nbn:de:gbv:ilm1-2016000052

Inhaltsverzeichnis

Abkürzungsverzeichnis	6
Abbildungsverzeichnis.....	7
Tabellenverzeichnis.....	8
1 Einleitung	10
1.1 Problemstellung.....	10
1.2 Zielsetzung.....	16
1.3 Methodik	17
1.3.1 Forschungsmethodische Einordnung der Arbeit	17
1.3.2 Methodische Vorgehensweise bei der Literaturanalyse.....	26
1.4 Aufbau der Arbeit.....	36
2 Information Lifecycle Management (ILM)	37
2.1 Begriff des Lebenszyklus.....	37
2.2 Begriff des ILM.....	38
2.3 Unstrukturierte Daten im Kontext von ILM	40
2.4 Anforderungen an ILM	42
2.4.1 Klassifizierung der Informationen im Verlauf ihres Lebenszyklus	42
2.4.2 Technische und funktionale Anforderungen	43
2.4.3 Rechtliche Anforderungen	48
2.4.3.1 Rechtliche Anforderungen an die Übertragung bzw. den Transfer der Informationen.....	49
2.4.3.2 Rechtliche Anforderungen an die Ablage und Aufbewahrung der Informationen	49
2.4.3.3 Rechtliche Anforderungen an die Wiedergabe der Informationen.....	51
2.4.3.4 Erstellung einer Verfahrensdokumentation	51
3 Cloud Computing.....	52
3.1 Begriff des Cloud Computings.....	52
3.2 Dimensionen des Cloud Computings.....	54
3.2.1 Organisatorische Dimension - Public, Private und Hybrid Cloud.....	54
3.2.2 Technische Dimension - Cloud-Computing-Dienste.....	55
3.3 Unstrukturierte Daten im Kontext des Cloud Computings.....	56
3.4 Anforderungen an das Cloud Computing	57
3.4.1 Technische und Funktionale Anforderungen	57

3.4.2	Rechtliche Anforderungen	60
3.4.2.1	Rechtliche Anforderungen an die Übertragung bzw. den Transfer der Informationen.....	61
3.4.2.2	Rechtliche Anforderungen an die Ablage und Aufbewahrung der Informationen	62
3.4.2.3	Rechtliche Anforderungen an die Wiedergabe der Informationen	65
3.4.2.4	Erstellung einer Verfahrensdokumentation	65
4	Wirtschaftlichkeitsanalyse	66
4.1	Begriff der Wirtschaftlichkeitsanalyse	66
4.2	Methoden der Wirtschaftlichkeitsrechnung	67
4.2.1	Ermittlungsmodelle.....	67
4.2.1.1	Statische Methoden.....	68
4.2.1.2	Dynamische Methoden	69
4.2.2	Quantitative Entscheidungsmodelle	72
4.2.2.1	Analytische Methoden	72
4.2.2.2	Simulationsmethoden	73
4.2.3	Qualitative Entscheidungsmodelle	75
4.3	Wirtschaftlichkeitsanalyse im ILM.....	76
5	Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM	79
5.1	Konstruktion eines Vorgehensmodells für das ILM	79
5.1.1	Vorstellung von ILM-Vorgehensmodellen	79
5.1.2	Konstruktion eines ILM-Vorgehensmodells.....	85
5.1.2.1	Analyse der IS-Architektur	91
5.1.2.2	Beschreibung relevanter Datenbestände.....	93
5.1.2.3	Klassifizierung der Daten	94
5.1.2.4	Verlagerung der Daten.....	96
5.1.3	Automatisierte Informationsklassifizierung im Rahmen des ILM	97
5.1.3.1	Vorstellung und Vergleich von Konzepten zur automatisierten Klassifizierung unstrukturierter Daten	97
5.1.3.1.1	Vorstellung der Konzepte	97
5.1.3.1.2	Vergleich der vorgestellten Konzepte.....	101

5.1.3.2 Anforderungen an Konzepte zur automatisierten Klassifizierung unstrukturierter Daten	104
5.1.3.2.1 Verwendung des Nutzungsgrads als Klassifizierungskriterium..	104
5.1.3.2.2 Implementierung geeigneter Klassifizierungsfunktionen.....	105
5.1.3.2.3 Berücksichtigung des Anwender- und Administratorenwissens.	105
5.1.3.2.4 Berücksichtigung rechtlicher Rahmenbedingungen	106
5.1.3.2.5 Kostenreduktion und Wirtschaftlichkeit	106
5.1.3.2.6 Berücksichtigung großer Datenmengen	107
5.2 Konstruktion und Implementierung eines Verfahrens zur Wirtschaftlichkeits- analyse der automatisierten Verwaltung unstrukturierter Daten im ILM	108
5.2.1 Anforderungen an die Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM.....	108
5.2.2 Konstruktion eines Verfahrens zur Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM	111
5.2.2.1 Klassifizierung der Daten	114
5.2.2.1.1 Festlegung von Klassifizierungskriterien	115
5.2.2.1.2 Festlegung von Messmethoden und -größen.....	116
5.2.2.1.3 Erhebung von Messwerten	117
5.2.2.1.3.1 Lebenszyklen der Daten	117
5.2.2.1.3.1.1 Sinkender Nutzungsgrad	119
5.2.2.1.3.1.2 Steigender Nutzungsgrad	121
5.2.2.1.3.1.3 Konstanter Nutzungsgrad	122
5.2.2.1.3.1.4 Saisonaler Nutzungsgrad	123
5.2.2.1.3.1.5 Sinkender Trend-Saisonaler Nutzungsgrad.....	123
5.2.2.1.3.1.6 Steigender Trend-Saisonaler Nutzungsgrad.....	125
5.2.3 Implementierung eines Verfahrens zur Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM	126
5.2.3.1 Klassifizierungsfunktionen.....	126
5.2.3.1.1 Implementierung der Klassifizierungsfunktionen.....	126
5.2.3.1.1.1 Szenarien zur Klassifizierung unstrukturierter Daten	129
5.2.3.1.1.1.1 Szenario 1: Klassifizierung der Daten ohne Cloud-Ebene.....	129
5.2.3.1.1.1.2 Szenario 2: Klassifizierung der Daten mit Cloud-Ebene .	131
5.2.3.2 Implementierung der Verlagerungsfunktionen	133

5.2.3.3 Kostenfunktionen	135
5.2.3.3.1 Kostenarten und Kostenmodell	135
5.2.3.3.1.1 Anschaffungskosten	139
5.2.3.3.1.2 Betriebskosten	140
5.2.3.3.1.3 Sonstige Kosten	143
5.2.3.3.2 Implementierung der Kostenfunktionen	143
5.2.3.3.2.1 Szenarien zur Klassifizierung unstrukturierter Daten	143
5.2.3.3.2.1.1 Szenario 1: Klassifizierung der Daten ohne Cloud-Ebene	143
5.2.3.3.2.1.1.1 Kostenfunktion zur Berechnung der Speicherplatzkosten	143
5.2.3.3.2.1.1.2 Kostenfunktion zur Berechnung der Administrationskosten	144
5.2.3.3.2.1.1.3 Kostenfunktion zur Berechnung der Verlagerungskosten	145
5.2.3.3.2.1.2 Szenario 2: Klassifizierung der Daten mit Cloud-Ebene ..	145
5.2.3.3.2.1.2.1 Kostenfunktion zur Berechnung der Speicherplatzkosten	145
5.2.3.3.2.1.2.2 Kostenfunktion zur Berechnung der Administrationskosten	146
5.2.3.3.2.1.2.3 Kostenfunktion zur Berechnung der Verlagerungskosten	147
5.2.3.3.3 Interpretation und Ergänzung der Messwerte	151
5.3 Demonstration der Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM	152
5.3.1 Sinkender Nutzungsgrad	155
5.3.2 Steigender Nutzungsgrad	160
5.3.3 Konstanter Nutzungsgrad	163
5.3.4 Saisonaler Nutzungsgrad	166
5.3.5 Sinkender Trend-Saisonaler Nutzungsgrad	168
5.3.6 Steigender Trend-Saisonaler Nutzungsgrad	171
5.3.7 Zusammenfassung der Wirtschaftlichkeitsanalyse	173
5.4 Evaluation der Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM	175

6	Schlussbetrachtungen.....	179
6.1	Zusammenfassung und kritische Würdigung der Arbeit	179
6.2	Ausblick	185
	Literaturverzeichnis.....	188
Anhang 1	Datenmengen und Wachstum der Daten	242
Anhang 2	Nutzungsgrade ohne Rauschfaktor	243
Anhang 3	Flussdiagramme für Klassifizierungsfunktionen.....	244
Anhang 3.1	Klassifizierungsfunktion K_1 für Szenario 1	244
Anhang 3.2	Klassifizierungsfunktion K_2 für Szenario 2	245
Anhang 4	Implementierung der Klassen und Klassifizierungsfunktionen.....	246
Anhang 4.1	Implementierung für Szenario 1	246
Anhang 4.2	Implementierung für Szenario 2	247
Anhang 5	Kostenarten und Kostenmodell	248
Anhang 5.1	Anschaffungskosten	248
Anhang 5.2	Betriebskosten - Parameter der Kostenfunktionen	249
Anhang 5.3	Betriebskosten und Betriebskostenfaktoren im ILM.....	250

Abkürzungsverzeichnis

BITKOM	Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V.
CRM	Customer Relationship Management
DSRM	Design Science Research Methodology
ERP	Enterprise Resource Planning
HSM	Hierarchical Storage Management
ILM	Information Lifecycle Management
IaaS	Infrastructure as a Service
IS	Informationssystem
IT	Informationstechnologie
NIST	National Institute of Standards and Technology
NLS	Nearline Storage
PaaS	Platform as a Service
SaaS	Software as a Service
SNIA	Storage Networking Industry Association

SRM	Storage Resource Management
TCO	Total Cost of Ownership
WI	Wirtschaftsinformatik
XAM	eXtensible Access Method

Abbildungsverzeichnis

Abb. 1-1:	Phasen des Forschungsprozesses	18
Abb. 2-1:	Lebenszyklus von Informationen	38
Abb. 2-2:	Speicherhierarchie von ILM	43
Abb. 2-3:	Verwaltung der Daten im Verlauf ihres Lebenszyklus	46
Abb. 3-1:	Technische Dimensionen und Zielgruppen des Cloud Computings	55
Abb. 3-2:	Cloud-Computing-Architektur	58
Abb. 5-1:	ILM-Vorgehensmodell BITKOM	80
Abb. 5-2:	ILM-Vorgehensmodell StorageTek	82
Abb. 5-3:	ILM-Vorgehensmodell SNIA	83
Abb. 5-4:	ILM-Vorgehensmodell EMC	84
Abb. 5-5:	Vorgehensmodell des ILM	91
Abb. 5-6:	Vergleich der Konzepte	102
Abb. 5-7:	Klassifizierung der Daten	111
Abb. 5-8:	Aufbau und Ablauf der Wirtschaftlichkeitsanalyse	113
Abb. 5-9:	Sinkender Nutzungsgrad ohne und mit Rauschfaktor	119
Abb. 5-10:	Steigender Nutzungsgrad ohne und mit Rauschfaktor	121
Abb. 5-11:	Konstanter Nutzungsgrad ohne und mit Rauschfaktor	122
Abb. 5-12:	Saisonaler Nutzungsgrad ohne und mit Rauschfaktor	123
Abb. 5-13:	Sinkender Trend-Saisonaler Nutzungsgrad ohne und mit Rauschfaktor	124
Abb. 5-14:	Steigender Trend-Saisonaler Nutzungsgrad ohne und mit Rauschfaktor	125
Abb. A1-1:	Datenmengen und Wachstum der Daten	242
Abb. A2-1:	Sinkender, steigender und konstanter Nutzungsgrad (ohne Rauschfaktor)	243
Abb. A2-2:	Saisonaler, sinkender trend-saisonaler und steigender trend-saisonaler Nutzungsgrad (ohne Rauschfaktor)	243
Abb. A3-1:	Flussdiagramm für Klassifizierungsfunktion 1	244
Abb. A3-2:	Flussdiagramm für Klassifizierungsfunktion 2	245

Abb. A4-1: Klassen und Intervalle Szenario 1	246
Abb. A4-2: Nutzungsgrade und Klassifizierungsfunktion Szenario 1	247
Abb. A4-3: Klassen und Intervalle Szenario 2	247
Abb. A4-4: Nutzungsgrade und Klassifizierungsfunktion Szenario 2	248

Tabellenverzeichnis

Tab. 1-1: Forschungsprozess und Forschungsmethoden	25
Tab. 1-2: Schlagworte Literaturanalyse	27
Tab. 1-3: Literaturanalyse zum ILM (Zeitschriften und Konferenzen)	30
Tab. 1-4: Dokumentation Literaturanalyse ILM (Zeitschriften und Konferenzen) ...	33
Tab. 5-1: Kostenmodell und Kostenarten	138
Tab. 5-2: Verwaltung Daten ohne ILM, 5 Jahre	153
Tab. 5-3: Sinkender Nutzungsgrad, ILM ohne Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad	155
Tab. 5-4: Sinkender Nutzungsgrad, ILM mit Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad	158
Tab. 5-5: Steigender Nutzungsgrad, ILM ohne Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad	160
Tab. 5-6: Steigender Nutzungsgrad, ILM mit Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad	162
Tab. 5-7: Konstanter Nutzungsgrad, ILM ohne Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad	164
Tab. 5-8: Konstanter Nutzungsgrad, ILM mit Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad	165
Tab. 5-9: Saisonaler Nutzungsgrad, ILM ohne Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad	166
Tab. 5-10: Saisonaler Nutzungsgrad, ILM mit Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad	167
Tab. 5-11: Sinkender Trend-Saisonaler Nutzungsgrad, ILM ohne Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad	169
Tab. 5-12: Sinkender Trend-Saisonaler Nutzungsgrad, ILM mit Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad	170

Tab. 5-13: Steigender Trend-Saisonaler Nutzungsgrad, ILM ohne Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad.....	172
Tab. 5-14: Steigender Trend-Saisonaler Nutzungsgrad, ILM mit Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad.....	173
Tab. A5-1: Parameter der Kostenfunktionen.....	250
Tab. A5-2: Betriebskosten Speicherplatz (SBK)	251
Tab. A5-3: Betriebskostenfaktoren (BKF)	253
Tab. A5-4: Betriebskosten Administration (ABK) und allgemeine Betriebskosten (GBK).....	254
Tab. A5-5: Kosten Zugriffe Cloud.....	255
Tab. A5-6: Betriebskosten Verlagerung Daten (VBK) Speichermedien.....	255
Tab. A5-7: Kosten Anlegen von Dateien	256
Tab. A5-8: Kosten Datenexport Cloud	256
Tab. A5-9: Kosten Versand Speichermedien (10 bis 28,56 TB)	257
Tab. A5-10: Kosten Versand Speichermedien (100 bis 285,61TB)	257
Tab. A5-11: Betriebskostenfaktor ohne ILM	257
Tab. A5-12: Betriebskosten Administration (ABK) und allgemeine Betriebskosten (GBK) ohne ILM	258

1 Einleitung

1.1 Problemstellung

Information Lifecycle Management (ILM) zielt darauf ab, Informationen¹ zu klassifizieren und kostengünstig zu verwalten.² Informationen, auf die in einem Unternehmen oft zugegriffen wird, können von Informationen mit geringerer Zugriffshäufigkeit getrennt³ und den Unternehmensanforderungen entsprechend auf dem jeweils sinnvollsten Speichermedium bereitgestellt und verwaltet werden.⁴ Die wesentlichen Gründe für die Entstehung von ILM sind das rasche Datenwachstum⁵ und die zahlreichen rechtlichen Anforderungen.⁶ Das Datenwachstum der in Unternehmen vorhandenen Anwendungssysteme⁷ führt zu hohen Speicher-, Administrations- und Betriebskosten.⁸ Die Kosten für den Betrieb und die Administration der Datenträger bzw. Datenspeicher, wie beispielsweise die Datensicherung und die Aufbewahrung der Daten, sind um das vier- bis achtfache höher als die Beschaffungskosten der Datenträger.⁹ Die Daten müssen aufgrund der gesetzlichen Anforderungen für längere Zeiträume sicher aufbewahrt werden,¹⁰ was zu einer zusätzlichen Erhöhung der Kosten führt. Der Begriff ILM stammt ursprünglich

¹ Tatsächlich werden im Rahmen des ILM Daten, Dateien bzw. Dokumente analysiert und klassifiziert. Da in den meisten Publikationen der Betrachtungsgegenstand jedoch als Information bezeichnet wird, wird dieser Begriff auch in dieser Arbeit verwendet. In der Literatur wird anstelle „Information Lifecycle Management“ auch der Begriff „Data Lifecycle Management“ verwendet.

² Vgl. Maier, Hädrich, Peinl / Knowledge Infrastructures / 254-257; Thome, Sollbach / Modelle des ILM / 22-30. Zur Verwaltung von Daten sind Vorgehensweisen, Verfahren und Methoden für die Speicherung und das Wiederauffinden der Daten auf entsprechenden Speichermedien erforderlich (Vgl. hierzu Lassmann / Wirtschaftsinformatik / 218-219). Unter der Verwaltung der Informationen ist die Anschaffung und Administration von Speichermedien sowie die Speicherung, Klassifizierung und Verlagerung von Informationen zu verstehen.

³ Vgl. Moody, Walsh / Measuring Information / 496-512; Short / ILM / 2-37

⁴ Vgl. Abd-El-Malek et al. / Cluster-based storage / 59-72; Bhagwan et al. / Management of Data Storage / ; Glazer / Measuring Information / 99-110; Maier, Hädrich, Peinl / Knowledge Infrastructures / 109, 116; Mont / Privacy-aware ILM / 405-415; Petrocelli / Information Lifecycle Management / 180-181

⁵ Vgl. Gantz et al. / Exploding Universe / 2-5; Kanakamedala, Kaplan, Srinivasaraghavan / Data storage / 1-4; Zadok et al. / Reducing Storage Management Costs / 101-105

⁶ Vgl. Johnson, Agrawal / Law and Technology / ; Kaarst-Brown, Kelly / IT-Governance / ; Short / ILM / 2-37

⁷ Beispiele hierfür sind Dokumentenmanagement-, Data-Warehouse- oder ERP-Systeme.

⁸ Vgl. Abd-El-Malek et al. / self-storage / 55-62; Allen / waste storage dollars / ; Thome, Sollbach / Modelle des ILM / 22-30

⁹ Vgl. Glazer / Measuring Information / 99-110; Kaplan, Roy, Srinivasaraghavan / Demand data storage / 1; Forrest, Kaplan, Kindler / Data centers / 1-10; Forrest, Kaplan, Kindler / Energy Efficiency / 1-13; Gantz, Reinsel / Extracting Value / 4; Schadler / Email in the Cloud / 4

¹⁰ Vgl. Short / ILM / 2-37

aus dem nicht-akademischen Umfeld und wurde vorwiegend von Speicherherstellern, Beratungshäusern und Analysten geprägt. Zum Konzept des ILM liegen bislang nur wenige wissenschaftliche Untersuchungen vor.¹¹ Neben der Analyse der Wirtschaftlichkeit sind rechtliche und technische Fragen die wesentlichen Untersuchungsbereiche des Forschungsbereichs ILM. Die Analyse des Kosten- und Nutzenverhältnisses, das bei der Implementierung und Anwendung von ILM entsteht, wird in der Literatur bislang vernachlässigt.¹² Die Identifizierung relevanter Gesetze, gesetzesähnlicher Vorschriften und unternehmensinterner Regelungen¹³ bezüglich der Verwaltung von Informationen ist im ILM ebenfalls von Bedeutung.¹⁴ Aus technischer Sicht stehen die Weiterentwicklung von Speichersystemen und leistungsfähiger Hard- und Software¹⁵ im Mittelpunkt der Betrachtung. Vorgehensmodelle für ILM haben noch nicht den Stand erreicht, sich in der Praxis durchzusetzen.¹⁶ Die Gründe hierfür sind einerseits die gesetzlichen und unternehmensspezifischen Anforderungen an ILM, wie Anforderungen an die Aufbewahrung der Informationen¹⁷ sowie die Gewohnheiten der Anwender¹⁸ hinsichtlich der Nutzung der Informationen.¹⁹ Andererseits sind es technische Anforderungen, die eine praktikable Umsetzung von ILM erschweren. Als schwierig erweist sich insbesondere

¹¹ Vgl. Kaiser, Smolnik, Riempff / ILM-Framework / 483-484; Turczyk / ILM / 11-14; Turczyk / Wertzuweisung / 459-460

¹² Vgl. hierzu die Ergebnisse der Literaturanalyse Kapitel 1.3.2; Chen / Information valuation / 145-146; Verma et al. / Very Large File Systems / 166-167; Turczyk / ILM / 117-118. Die Verfasser von ILM-Konzepten konzentrieren sich auf die Entwicklung von Klassifizierungsfunktionen, um den Informationen entsprechend ihres Nutzungsgrads einen Klassifizierungswert zuzuweisen.

¹³ Die Analyse des Nutzungsverhaltens von Informationen steht dabei im Vordergrund.

¹⁴ Vgl. Bernard / ILM-Security / 26-30; Kaiser, Smolnik, Riempff / ILM-Framework / 483-494; Kaiser, Smolnik, Riempff / Verbesserte Compliance ILM / 30-38; Short / ILM / 2-37; Turczyk / Organisation ILM / 371-372

¹⁵ Vgl. Gillet, Mendel / Organic IT / 483-502

¹⁶ Vgl. Kaiser Smolnik / ILM für Dokumente / 104-105; Thome, Sollbach / Modelle des ILM / 150; Kaiser, Smolnik, Riempff / ILM-Framework / 483-484; Steinbinder, Bedau, Löw / Datenbanktechnologie für das ILM / 107

¹⁷ Vgl. Dauen / Aufbewahrungspflichten / 3-122; Short / ILM / 2-37

¹⁸ Vgl. Mesnier et al. / File classification / 44-51; Shah et al. / Classification for ILM / 1-15

¹⁹ Anwender und Administratoren können wichtige Hinweise zur Verwendung der Informationen geben, beispielsweise auf Basis der aktuellen und zukünftigen Projekt- bzw. Auftragslage oder der Verwendung von Informationen in verschiedenen Abteilungen eines Unternehmens. Aus diesem Grund sollte das Wissen der Anwender und Administratoren über die Systeme und Informationen berücksichtigt werden. Vgl. hierzu Bhagwan et al. / Management of Data Storage /; Douglass et al. / Delete-Optimized Storage /; Patrascu et al. / Autonomic computing / 140-145.

die Klassifizierung von Informationen.²⁰ Hierfür werden zahlreiche Kriterien benötigt, wie beispielsweise die Anzahl der Lese- und Schreibzugriffe.²¹ Die zeitnahe Erhebung, kontinuierliche Aktualisierung und Auswertung zahlreicher Klassifizierungskriterien ist ohne automatisierte Unterstützung praktisch kaum möglich und wirtschaftlich nicht sinnvoll - und erweist sich zugleich bei großen Datenbeständen²² als problematisch.²³ Eine zeitnahe bzw. monatliche Klassifizierung ist im Sinne der Kostensenkung von besonderer Relevanz.²⁴ Dienstleister oder Anbieter von Speicherplatz berechnen die Speicherkosten und den damit verbundenen Administrationsaufwand in der Regel auf Basis des Durchschnitts aus verschiedenen Tagen eines Rechnungsmonats.²⁵ Zudem müssen in einem Unternehmen verschiedene Dienstgütevereinbarungen über die Verfügbarkeit und die Antwortzeiten von Informationssystemen eingehalten werden.²⁶ Erste Konzepte zur Unterstützung der automatisierten Klassifizierung entwickelten beispielsweise Chen, Verma et al. und Turczyk.²⁷ Das Ziel dieser Konzepte ist die Senkung der Administrations- und Speicherkosten. Die Verfasser vernachlässigen allerdings die Evaluation des Kosten- und Nutzenverhältnisses,²⁸ das hinsichtlich Implementierung und Anwendung der Konzepte entsteht.²⁹ Die Evaluation der Modelle und Konzepte sowie die Untersuchung der Wirtschaftlichkeit stellen jedoch einen wesentlichen Aspekt der WI-

²⁰ Vgl. Abd-El-Malek et al. / self-storage / 55-62; Bhagwan et al. / Management of Data Storage / ; Chen / Information valuation / 135-146; Mesnier et al. / File classification / 44-51; Thome, Sollbach / Modelle des ILM / 22-30; Yates-Mercer, Bawden / Managing the paradox / 19-29

²¹ Vgl. Chen / Information valuation / 135-146; Ellard et al. / File prediction / 1-14

²² Als große Datenmengen werden zurzeit solche bezeichnet, die ein physisches Volumen von mehr als einem Terabyte aufweisen.

²³ Vgl. Bhagwan et al. / Management of Data Storage / ; Shah et al. / Classification for ILM / 1-15

²⁴ Vgl. Kapitel 5.1.3.1

²⁵ Vgl. Turczyk / Wertzuweisung / 462

²⁶ Vgl. Hiles / Service Level Agreements / 5-17

²⁷ Chen / Information valuation / 135-146; Turczyk / ILM / 11-14; Verma et al. / Very Large File Systems / 160-168; Vgl. Kapitel 3 zur Vorstellung und zum Vergleich weiterer Konzepte.

²⁸ Die Verfasser machen ebenfalls keine Aussagen darüber, mit welchen Kosten zu rechnen ist, die bei der praktischen Umsetzung der jeweiligen ILM-Konzepte entstehen. Vgl. hierzu Venable, Pries-Heje, Baskerville / Framework for Evaluation / 427.

²⁹ Vgl. Kapitel 5.1.3.2; Chen / Information valuation / 145-146; Verma et al. / Very Large File Systems / 166-167; Turczyk / ILM / 117-118. Die automatisierte Erhebung, Aktualisierung und Auswertung der Klassifizierungskriterien wird in den Konzepten vereinfacht, indem nicht jedes einzelne Datenobjekt zur Klassifizierung herangezogen wird. Es werden Anteile von Informationen innerhalb von Dateiordnern klassifiziert oder Stichproben bzw. kleine Datenbestände aus dem gesamten Datenbestand gezogen (Vgl. hierzu Shah et al. / Classification for ILM / 9-14; Turczyk / ILM / 90, 101-114, 140; Verma et al. / Very Large File Systems / 160-168). Durch diese Vorgehensweise wird die Leistungsfähigkeit der Informationssysteme positiv beeinflusst.

bzw. IS-Forschung dar.³⁰ Weiterhin ist zu beachten, dass der Einsatz von ILM nur bei großen und/oder schnell wachsenden Datenbeständen wirtschaftlich sinnvoll ist, da bei kleinen Datenbeständen der Aufwand für die Klassifizierung der Daten in keinem angemessenen Verhältnis zur Kostensenkung steht.³¹

Mit Hilfe von Cloud Computing werden IT-Dienstleistungen bedarfsgerecht und flexibel als Dienst über das Internet oder über ein Firmennetzwerk bereitgestellt.³² Das Cloud Computing ist eine Erweiterung des sogenannten Grid Computings.³³ Aufgrund der Nachfrage nach mehr Rechenleistung wurden Hochleistungsrechner miteinander verbunden, um durch die gemeinsame Nutzung von Ressourcen große Datenmengen verwalten und evaluieren zu können. Das Grid Computing wurde für den Einsatz in wissenschaftlichen Anwendungen entwickelt.³⁴ Für den kommerziellen Einsatz zur Verwaltung von Daten wurde der Begriff des Cloud Computings geprägt.³⁵ Anbieter von Cloud-Computing-Diensten stellen beispielsweise IT-Infrastrukturen³⁶ über das Internet zur Verfügung. Es werden Ressourcen wie Hardware, Netzwerk und Speicherkapazitäten bereitgestellt und nutzungsabhängig abgerechnet.³⁷ Gesetzliche und technische Fragen sind die wesentlichen Untersuchungsbereiche im Cloud Computing.³⁸ Aus technischer Sicht steht die Entwicklung von Schnittstellen für die Kommunikation und den Datenaustausch zwischen verschiedenen Cloud-Computing-Diensten im Vorder-

³⁰ Vgl. hierzu Venable, Pries-Heje, Baskerville / Framework for Evaluation / 427; Peffers et al. / Research Evaluation / 398-399. Die Untersuchung der Wirtschaftlichkeit wurde im ILM bislang vernachlässigt. In dieser Arbeit wird untersucht, welche Kosten durch die automatisierte Verwaltung der Daten im ILM verursacht werden und welcher Nutzen hierbei entsteht (Vgl. Kapitel 5.1.3.2 und Kapitel 5.3).

³¹ Vgl. Heinrich, Stelzer / Informationsmanagement / 275

³² Vgl. Dunkel et al./ Verteilte Anwendungen / 270; Baun et al. / Cloud Computing / 2

³³ Vgl. Dunkel et al./ Verteilte Anwendungen / 270; Vgl. Foster, Kesselmann, Tuecke / Anatomy of the Grid / 213; Vouk / Cloud Computing / 242-243

³⁴ Vgl. Fey / Grid-Computing / 4-6

³⁵ Technologisch gesehen besteht der wesentliche Unterschied zwischen Grid Computing und Cloud Computing darin, dass im Cloud Computing der Zugriff auf die Ressourcen mit Hilfe von Internet-basierten Anwendungen erfolgt (Vgl. hierzu Lawton / Moving OS to the Web / 16-19). Aus betriebswirtschaftlicher Sicht wurden für das Cloud Computing zusätzlich Geschäftsmodelle entwickelt, um den kommerziellen Einsatz der Technologie in Unternehmen zu ermöglichen (Vgl. Weiss / Computing in the Clouds / 16-25; Chang, Wills, De Roure / Cloud Business Models / 13-16).

³⁶ Unter dem Begriff IT-Infrastrukturen ist eine Ansammlung von Computern, Peripheriegeräten, Speichern und Netzwerkkomponenten zu verstehen (Vgl. Rumpel / IT-Infrastrukturen / 11).

³⁷ Vgl. Baun et al. / Cloud Computing / 18

³⁸ Vgl. Kapitel 3.4

grund.³⁹ Aus rechtlicher Sicht stehen der sichere Transfer und die sichere Speicherung der Daten im Mittelpunkt der Betrachtung.⁴⁰ Bislang existieren wenige wissenschaftliche Untersuchungen dazu, wie das Cloud Computing im ILM eingesetzt werden kann, um unstrukturierte Daten zu verwalten.⁴¹ Mit Hilfe von ILM können Informationen entsprechend den Unternehmensanforderungen auf dem jeweils sinnvollsten Speichermedium bereitgestellt und verwaltet werden.⁴² Das Konzept des Cloud Computings steht ebenfalls für die kostengünstige Verwaltung unstrukturierter Daten.⁴³ Die Nutzung des Cloud Computings als Dienst über das Internet kann beispielsweise kostengünstiger sein, als die Verwaltung der Informationen auf unternehmenseigenen Speichermedien im ILM.⁴⁴ Aufgrund rechtlicher Vorschriften⁴⁵ können allerdings nicht sämtliche Daten

³⁹ Vgl. Weinhardt et al. / Cloud-Computing / 460; Cheptsov et al. / Scalable Cloud Platform / 8-12

⁴⁰ Vgl. Rittinghouse, Ransome / Cloud Computing / 135; Sosinsky / Cloud Computing Bible / 398; Vgl. Kapitel 3.4.2. Da im Cloud Computing auch Dienste über das Internet bereitgestellt und genutzt werden, ist neben den rechtlichen Anforderungen die IT-Sicherheit von besonderer Relevanz. Sicherheit im Sinne der Informationsverarbeitung ist dabei ein Zustand, in dem alle sicherungswürdigen Informationen vor möglichen Beeinträchtigungen bewahrt sind (Vgl. Stelzer / Sicherheitsstrategien / 23, Kapitel 2.4.3). Solche Beeinträchtigungen sind beispielsweise menschliches und technisches Versagen, Fehler bei der Informationsübertragung, Spionage oder höhere Gewalt (Vgl. BSI / Sicherheitshandbuch /). Insbesondere Beeinträchtigungen durch Spionage erhalten nach der Aufdeckung des weltweiten Abhörskandals durch die National Security Agency (NSA) einen besonderen Stellenwert (Vgl. Venhaus, Haselbeck, Wintermann / Schatten der Überwachung / 40-50; Wicker / Death of Privacy / 129-134).

⁴¹ Vgl. Du / Optimal Storage / 730-746; Sotomayor et al. / Private and Hybrid Clouds / 14-21; Jo et al. / Virtual Disk / 375-377; Hill et al. / Guide Cloud Computing / 250-251; Kapitel 1.3.2

⁴² Vgl. Abd-El-Malek et al. / Cluster-based storage / 59-72; Bhagwan et al. / Management of Data Storage / ; Glazer / Measuring Information / 99-110; Maier, Hädrich, Peinl / Knowledge Infrastructures / 109, 116; Mont / Privacy-aware ILM / 405-415; Petrocelli / Information Lifecycle Management / 180-181

⁴³ Vgl. Gantz, Reinsel / Extracting Value / 2, 4, 11; Sriram / Cloud-Scale / 381-390; Mugg / Storage für KMU / 33. Auch die In-Memory-Technologie stellt eine Möglichkeit zur Verwaltung von Informationen dar. Es erfolgt die Speicherung der Informationen im Arbeitsspeicher (Vgl. hierzu Loos et al. / In-Memory / 384-385; Garcia-Molina, Salem / Main Memory Database / 509-516). In-Memory-Technologien kommen bei der Analyse strukturierter Daten zum Einsatz (Vgl. hierzu Plattner, Zeier / In-Memory Data Management / 2, 7, 22). Die Analyse und Auswertung der Informationen wird somit beschleunigt (Vgl. hierzu Plattner, Zeier / In-Memory Data Management / 3, 9; Plattner / Database Approach / 1-2). Die Datenhaltung im Arbeitsspeicher ist nur mit sogenannten Snapshots, Replikationen oder Transaction Logs möglich, für eine dauerhafte Speicherung von unstrukturierten Daten, wie Dokumenten oder E-Mails, ist die In-Memory-Technologie nicht geeignet (Vgl. hierzu Schütte / Einsatzpotential In-Memory / 1-12; Oates et al. / Spring und Hibernate / 32-33; Shirazi / Performance Tuning / 484-487; Mall / Real-Time Systems / 216-217). Aus diesem Grund wird die In-Memory-Technologie in dieser Arbeit nicht weiter betrachtet.

⁴⁴ Wenn Informationen im Cloud Computing als Dienst über das Internet verwaltet werden, können die Kosten für die Administration unternehmenseigener Speichermedien weiter gesenkt werden (vgl. Schadler / Email in the Cloud / 3-5).

⁴⁵ Die Betrachtung rechtlicher Vorschriften beim Einsatz von ILM-Konzepten und der Verwaltung von Informationen im Cloud Computing wird in der Forschung bislang vernachlässigt (Vgl. Mont / Identity Lifecycle Management / 397-426; Gantz, Reinsel / Extracting Value / 9-10). Der Fokus bisheriger Publikationen liegt vorwiegend auf der technischen Weiterentwicklung dieser Konzepte. In dieser Arbeit werden die relevanten rechtlichen Vorschriften zur Verwaltung von Informationen im Cloud Computing

mit Hilfe von Diensten über das Internet verwaltet werden.⁴⁶ Cloud Computing kann jedoch eine sinnvolle Erweiterung des ILM darstellen, indem nicht nur die Klassifizierung und Verwaltung von Informationen auf unternehmenseigenen Speichermedien betrachtet wird, sondern auch die Verwaltung der Informationen mit Hilfe von Diensten über das Internet. Somit können zusätzliche Kostensenkungspotentiale entsprechend der Unternehmensanforderungen und zu verwaltenden Informationen genutzt werden.⁴⁷

Neben strukturierten Daten zählen insbesondere unstrukturierte Daten⁴⁸ zum Gegenstandsbereich von ILM und Cloud Computing.⁴⁹ Die vorliegende Arbeit beschäftigt sich mit der Verwaltung unstrukturierter Daten, da etwa achtzig Prozent der Daten in einem Unternehmen unstrukturierte Daten sind und die wirtschaftliche Verwaltung unstrukturierter Daten wissenschaftlich bislang wenig erforscht ist.⁵⁰ Die Kriterien zur Klassifizierung unstrukturierter Daten lassen sich mit geringem Aufwand bestimmen.⁵¹ Bei strukturierten Datensätzen ist dies aufgrund der in einer Datenbank vorkommenden starken Interdependenzen wesentlich schwieriger.⁵² Beispiele für schnell wachsende Datenbestände in Unternehmen sind E-Mails⁵³, Verbindungsdaten von Telekommunika-

erörtert.

⁴⁶ Aufgrund der rechtlichen Anforderungen an die Ablage und Aufbewahrung von Informationen muss für bestimmte Daten, wie beispielsweise vertrauliche und personenbezogene Daten, das Speichermedium und der exakte Aufbewahrungsort der Informationen bekannt sein (Vgl. hierzu Kapitel 2.4.3.2). Bei der Nutzung von Cloud-Computing-Diensten ist allerdings nicht immer feststellbar, wo sich die gespeicherten Daten befinden und ob der Dienste-Anbieter die datenrechtlichen Vorschriften einhält (Vgl. Kriesi / Rechtliche Aspekte von SaaS / 40-41; Gantz, Reinsel / Extracting Value / 9).

⁴⁷ Vgl. Baun et al. / Cloud Computing / 9

⁴⁸ Unstrukturierte Daten enthalten bezüglich ihrer Struktur keine Vorgaben zur Anordnung von Datenfeldern bzw. Attributen und deren Wertebereichen. Beispiele hierfür sind E-Mails, Textdokumente, Präsentationen, Videos oder Audio-Daten. (Vgl. hierzu Stahlknecht, Hasenkamp / Wirtschaftsinformatik / 135-137; Vgl. Kapitel 2.3).

⁴⁹ Vgl. Kanakamedala, Kaplan, Srinivasaraghavan / Data storage / 1-4

⁵⁰ Vgl. hierzu Gantz et al. / Digital Universe / 12; Schoop / Informationsmanagement / 556-568; Vouk / Cloud Computing / 242-243; McKinsey / Challenge Big Data / 2-3; Reichman, Whiteley, Chi / File Storage Costs / 2-5.

⁵¹ Durch die Klassifizierung der unstrukturierten Daten können diese entsprechend ihrer Klasse auf einem geeigneten Speichermedium verwaltet werden (Vgl. hierzu Kapitel 2.4.1 und 2.4.2). Vgl. Bhagwan et al. / Management of Data Storage /; Chandra, Gehani, Yu / Automated Storage / 12; Chen / Information valuation / 135-146; Douglass et al. / Delete-Optimized Storage /; Gavish, Sheng / Dynamic file migration / 177-189; Shah et al. / Classification for ILM / 1-15; Skyrme / Add value to business / 20-25.

⁵² Vgl. Ballou, Smith, Wang / Quality Estimation Database / 639-650; Chaudhuri, Weikum / Database System Architecture / 1-10; Kaiser, Smolnik, Riempp / ILM-Framework / 484; Vassiliadis et al. / Warehouse Process Management / 205-236; Kosler, Matthesius, Stelzer / Klassifizierung von Informationen / 129-145

⁵³ BITKOM / Leitfaden ILM / 4

tionsanbietern, Abrechnungsdaten von Energieversorgern, digitale Bildarchive von Krankenhäusern⁵⁴ oder Geschäftsdokumente von Banken und Versicherungsgesellschaften⁵⁵.⁵⁶ Etwa vierundneunzig Prozent des weltweiten Datenaufkommens sind unstrukturierte Daten.⁵⁷ Dies ist zurückzuführen auf das rasche Wachstum von Dokumenten, Bilddaten, Videodateien, Audiodateien und Profildaten in Unternehmen, privaten Haushalten und sozialen Netzwerken.⁵⁸

1.2 Zielsetzung

Ziel der Arbeit ist die Konstruktion und Implementierung eines Verfahrens zur Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM.⁵⁹ Da ILM auf die kostengünstige Verwaltung von Informationen abzielt,⁶⁰ wird untersucht, welche Kosten- und Nutzenfaktoren für das ILM von Relevanz sind.⁶¹ Wie die Wirtschaftlichkeit unter Berücksichtigung unterschiedlich großer Datenmengen, variierender Zugriffshäufigkeiten und verschiedener Speicherebenen gemessen und evaluiert werden kann, ist ebenfalls Gegenstand der Betrachtungen. Neben der Verwaltung unstrukturierter Daten auf unternehmenseigenen Speichermedien wird zusätzlich das Konzept des Cloud Computings⁶² zur Verwaltung der Daten mit Hilfe von Diensten über das Internet⁶³ betrachtet.

⁵⁴ Vgl. Heinrich, Stelzer / Informationsmanagement / 275-276

⁵⁵ Vgl. Turczyk / Organisation ILM / 371-372

⁵⁶ Zur Beschreibung der schnell wachsenden Bestände unstrukturierter Daten wird ebenfalls der Begriff Big Data verwendet. Vgl. hierzu Mohanty, Jagadeesh, Srivatsa / Big Data Imperatives / 151-152.

⁵⁷ Vgl. Gantz et al. / Exploding Universe / 2-5; Gantz, Reinsel / Extracting Value / 1-2

⁵⁸ Vgl. Kaplan, Roy, Srinivasaraghavan / Demand data storage / 1-3; Smullen et al. / Unstructured Data Processing Applications / 293; Du et al. / Optimal Storage / 731-735

⁵⁹ Vgl. Kapitel 5

⁶⁰ Vgl. Kapitel 1.1

⁶¹ Vgl. Kapitel 5.2.3.3

⁶² Vgl. Kapitel 3.1

⁶³ Vgl. Kapitel 3.2.1

1.3 Methodik

1.3.1 Forschungsmethodische Einordnung der Arbeit

Informationssysteme⁶⁴ sind Forschungsgegenstand der im angloamerikanischen Raum vorherrschenden Disziplin Information Systems (IS) und der im deutschsprachigen Raum vertretenen Disziplin Wirtschaftsinformatik (WI). Dennoch unterscheiden sich beide Disziplinen hinsichtlich ihrer Forschungsmethodik.⁶⁵ Das in der IS vorherrschende verhaltenswissenschaftliche Paradigma entstammt vorwiegend der empirischen Sozialforschung. Im Vordergrund steht die Untersuchung des Zusammenhangs zwischen organisatorischen und sozialen Phänomenen und der Gestaltung, Implementierung und Nutzung von Informations- und Kommunikationssystemen.⁶⁶ Im Gegensatz dazu herrscht in der WI das konstruktionswissenschaftliche Paradigma vor, das sich mit der Konstruktion und Evaluation von praxisrelevanten Artefakten bzw. Modellen, Methoden, Verfahren oder Systemen beschäftigt.⁶⁷ An Kritik wird dem verhaltenswissenschaftlichen Paradigma entgegengebracht, dass die Forschungsergebnisse für Entscheidungsträger in der Praxis von geringem Nutzen seien.⁶⁸ Demgegenüber werfen Kritiker dem konstruktionswissenschaftlichen Paradigma einen Mangel an wissenschaftlichem Anspruch vor, da Forschungsergebnisse zu stark auf die Praxisrelevanz abzielen.⁶⁹ Beide Disziplinen können sich in Abhängigkeit vom Forschungsgegenstand inhaltlich ergänzen.⁷⁰ Forschungsergebnisse über organisatorische oder soziale Phänomene können beispielsweise die Grundlage für die Konstruktion von Artefakten darstellen.⁷¹

Zur Umsetzung der in Kapitel 1.2 festgelegten Ziele wird in dieser Arbeit eine konstruktionswissenschaftliche Methodik gewählt. Peffers et al. beschreiben in ihrem Vorgehensmodell „Design Science Research Methodology“ (DSRM) eine Vorgehensweise

⁶⁴ Informations- und Kommunikationssysteme sind soziotechnische Systeme, die aus den Komponenten Mensch, Aufgabe und Technik bestehen. Vgl. hierzu Heinrich, Heinzl, Roithmayr / Wirtschaftsinformatik / 14-15.

⁶⁵ Vgl. Schauer, Frank / WI und IS / 127

⁶⁶ Vgl. Gregor / Theory of IS / 611-642; Hevner et al. / Design Science in IS / 75-105

⁶⁷ Vgl. Becker, Pfeiffer / Forschung in der WI / 1-17; Hevner et al. / Design Science in IS / 76; Simon / Artificial / 113-114

⁶⁸ Vgl. Frank / Research Methods in IS / 3-4

⁶⁹ Vgl. Frank / Research Methods in IS / 5-6

⁷⁰ Vgl. Becker, Pfeiffer / Konzeptuelle Modellierung / 11; Zelewski / Publikationspraxis / 88-89

⁷¹ Vgl. Frank / Research Methods in IS / 40-61

zur Bearbeitung konstruktionswissenschaftlicher Fragestellungen. In Abbildung 1-1 sind die sechs Phasen und mögliche Einstiegspunkte des Forschungsprozesses der konstruktionswissenschaftlichen Methodik abgebildet.

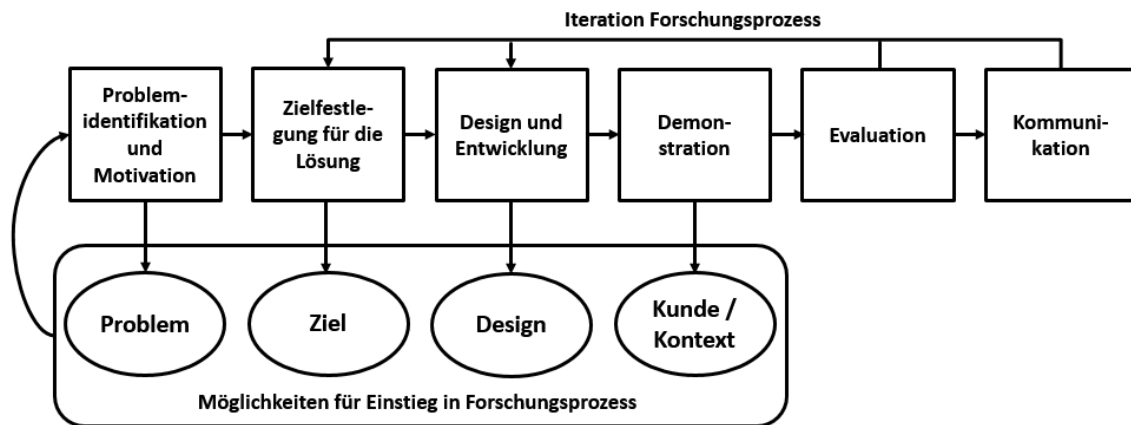


Abb. 1-1: Phasen des Forschungsprozesses⁷²

In der Phase „Problemidentifikation und Motivation“ wird das Forschungsproblem beschrieben. Der Forschungsprozess muss dabei relevante Probleme behandeln⁷³ und der Wertbeitrag einer zu entwickelnden Lösung muss ermittelt werden. Bei einer komplexen Problemstellung sollte das Problem in einzelne Teilprobleme aufgeteilt werden. Somit wird die Komplexität reduziert und es können in der Phase „Zielfestlegung für die Lösung“ Anforderungen an das Artefakt auf Basis der Teilprobleme abgeleitet werden.⁷⁴ Die Ziele des Artefakts werden festgelegt. Ziel eines zu entwickelnden Artefakts ist es beispielsweise, besser zu einer Problemlösung beizutragen, als bereits existierende Artefakte.⁷⁵ In der Phase „Design und Entwicklung“ erfolgen der Entwurf und die Entwicklung des Artefakts. Artefakte können Konzepte, Modelle, Methoden, Verfahren oder Systeme darstellen.⁷⁶ Bevor das Artefakt entwickelt wird, werden Funktionalität und Architektur des Artefakts entworfen. Die „Demonstration“ eignet sich dazu, die

⁷² In Anlehnung an Peffers et al. / Research Methodology / 54

⁷³ Beispielsweise aus dem Bereich Forschung und Entwicklung oder aus der Unternehmenspraxis. Vgl. hierzu Rossi, Sein / Design Research / 9-12; Hevner et al. / Design Science in IS / 75-105.

⁷⁴ Vgl. Peffers et al. / Research Methodology / 52

⁷⁵ Dies setzt voraus, dass bereits existierende Artefakte identifiziert und analysiert wurden. Weiterhin ist denkbar, dass ein Artefakt auf die Lösung bislang ungelöster Probleme abzielt bzw. auf Probleme, für die noch keine vergleichbaren Artefakte existieren. Vgl. Peffers et al. / Research Methodology / 55.

⁷⁶ Vgl. Järvinen / Action Research / 37-54; Hevner et al. / Design Science in IS / 75-105

Funktionsweise des Artefakts zu überprüfen und nachzuweisen.⁷⁷ Ziel der „Evaluation“ ist es, die zur Lösung eines bestimmten Problems definierten Objekte auf der Grundlage eines Systems von Kriterien zielbezogen zu beurteilen.⁷⁸ Die Demonstration selbst kann bereits als ein Teil der „Evaluation“ betrachtet werden, da durch den Einsatz von Simulationen, Experimenten oder Fallstudien zumindest teilweise aufgezeigt wird, wie und ob das Artefakt zur Lösung der definierten Problemstellung beiträgt.⁷⁹ Die Publikation der identifizierten Problemstellung und seiner Bedeutung sowie des zur Problemlösung entwickelten Artefakts und des damit verbundenen Nutzens erfolgt in der Phase „Kommunikation“.⁸⁰ Die sechs Phasen des Vorgehensmodells müssen bei der Bearbeitung von konstruktionswissenschaftlichen Fragestellungen vollständig durchlaufen werden.⁸¹ Aufgrund des iterativen Charakters des Vorgehensmodells können die einzelnen Phasen mehrfach durchlaufen werden,⁸² eine sequentielle Abarbeitung der Phasen ist jedoch nicht notwendig. Mit welcher Phase begonnen wird, wird über Einstiegspunkte festgelegt. Die definierte Problemstellung, die Zielsetzung, ein bereits entwickeltes Artefakt oder die vorliegenden Ergebnisse aus einem Kundenprojekt dienen als mögliche Einstiegspunkte in den Forschungsprozess.

Da sich das DSRM von Peffers et al. in der Wissenschaft etabliert hat und als allgemein anerkannt gilt,⁸³ bildet das Vorgehensmodell die Grundlage für den Forschungsprozess dieser Arbeit. Im Folgenden wird erörtert, welcher Einstiegspunkt gewählt wird, wie die einzelnen Phasen des Forschungsprozesses durchlaufen werden und welche Forschungsmethoden in den jeweiligen Phasen eingesetzt werden:

- Problemidentifikation und Motivation: Die Problemstellung⁸⁴ der vorliegenden Arbeit wird aus den Ergebnissen einer Literaturanalyse⁸⁵ zum aktuellen Stand

⁷⁷ Vgl. hierzu Peffers et al. / Research Methodology / 55; Walls, Widmeyer, El Sawy / Design Theory / 36-59.

⁷⁸ Vgl. Heinrich / Evaluationsforschung / 9; Mertens / Evaluation / 47

⁷⁹ Vgl. Venable, Pries-Heje, Baskerville / Framework for Evaluation / 428

⁸⁰ Vgl. Archer / Method for designers / 57-82; Hevner et al. / Design Science in IS / 75-105

⁸¹ Vgl. Peffers et al. / Research Methodology / 56

⁸² Wenn bspw. aufgrund der Evaluationsergebnisse der Bedarf zur Anpassung des Artefakts ermittelt wird, so kann die Phase „Design und Entwicklung“ erneut durchlaufen werden, um ggf. die Funktionalität des Artefakts zu erweitern.

⁸³ Vgl. Hevner, Chatterjee / Design Research / 28; Falk, Griesberger, Leist / Artifact / 89

⁸⁴ Vgl. Abbildung 1-1, Phasen „Problemidentifizierung und Motivation“

der Forschung zur Wirtschaftlichkeitsanalyse der automatisierten Klassifizierung unstrukturierter Daten im ILM abgeleitet.⁸⁶ Wie bei der Literaturanalyse methodisch vorgegangen wurde, wird in Kapitel 1.3.2 erörtert. Das Wachstum der Unternehmensdaten führt zu hohen Speicher-, Administrations- und Betriebskosten. Steigende Speicher- und Administrationskosten in Unternehmen, hervorgerufen durch das Wachstum der Unternehmensdaten, stellen ein relevantes Problem für Unternehmen und Organisationen dar.⁸⁷ Durch die Verwaltung der Daten im ILM können diese Kosten gesenkt werden. Die Analyse der Wirtschaftlichkeit im ILM wurde in der Literatur bislang vernachlässigt.⁸⁸ Der Einstiegspunkt in den Forschungsprozess dieser Arbeit ist problemorientiert.⁸⁹ Die zentralen Forschungsfragen dieser Arbeit lauten: Inwiefern kann die automatisierte Klassifizierung unstrukturierter Daten im ILM zu einer Kostensenkung in Unternehmen beitragen und welche Kosten- und Nutzenaspekte sind dabei von Relevanz? Wie kann die Wirtschaftlichkeit im ILM, unter Berücksichtigung verschiedener Lebenszyklen und Datenmengen, gemessen und evaluiert werden? Weiterhin ist von Interesse, inwieweit die Nutzung von Cloud-Computing-Diensten im ILM zu einer Veränderung der Wirtschaftlichkeit beiträgt.

- Zielfestlegung für die Lösung: Aus der Problemstellung wird die Zielsetzung der Arbeit abgeleitet.⁹⁰ Ziel der Arbeit ist die Konstruktion und Implementierung eines Verfahrens zur Wirtschaftlichkeitsanalyse der automatisierten Klassifizierung unstrukturierter Daten im ILM. Anforderungen an das zu entwickelnde Verfahren resultieren aus der zuvor durchgeführten Literaturanalyse.⁹¹ Basierend auf technischen und rechtlichen Anforderungen sowie Anforderungen an die automatisierte Klassifizierung werden Anforderungen an die Wirtschaftlichkeits-

⁸⁵ Vgl. Kapitel 1.3.2; Kapitel 2 bis Kapitel 4

⁸⁶ Vgl. Kapitel 1.1, Kapitel 1.2, Kapitel 4.3 und Kapitel 6.1

⁸⁷ Vgl. Abd-El-Malek et al. / self-storage / 55-62; Allen / waste storage dollars / ; Thome, Sollbach / Modelle des ILM / 22-30. Vgl. hierzu die Anforderung der Problemrelevanz für den Forschungsprozess Abbildung 1-1.

⁸⁸ Vgl. hierzu Chen / Information valuation / 145-146; Verma et al. / Very Large File Systems / 166-167; Turczyk / ILM / 117-118

⁸⁹ Vgl. hierzu Abbildung 1-1, „Möglichkeiten für Einstieg in den Forschungsprozess“

⁹⁰ Vgl. Abbildung 1-1, Phase „Zielfestlegung für die Lösung“

⁹¹ Vgl. Peffers et al. / Research Methodology / 55 zur Ableitung von Anforderungen an das zu erstellende Artefakt bzw. Verfahren auf Basis einer Literaturanalyse; Vgl. Kapitel 5.2 und Kapitel 5.3.

analyse der automatisierten Klassifizierung⁹² unstrukturierter Daten im ILM definiert.⁹³

- Design und Entwicklung: Für die Zielerreichung wird ein Artefakt bzw. Verfahren konstruiert⁹⁴ und prototypisch implementiert. Um die Wirtschaftlichkeitsanalyse im ILM durchführen zu können, müssen zunächst die wesentlichen Teilaufgaben des ILM beschrieben werden. Auf der Basis von publizierten ILM-Vorgehensmodellen wird hierzu ein Vorgehensmodell⁹⁵ entwickelt. Wie die einzelnen Phasen des Vorgehensmodells hergeleitet und entwickelt wurden, wird in Kapitel 5.1.2 erörtert. Im Gegensatz zu bereits existierenden Vorgehensmodellen⁹⁶ berücksichtigt das Vorgehensmodell neben der Verwaltung unternehmensinterner Speichermedien zusätzlich das Konzept des Cloud Computings.⁹⁷ Die Phase „Klassifizierung der Daten“ des ILM-Vorgehensmodells bildet die konzeptionelle Basis für die Konstruktion und Implementierung des Verfahrens zur Wirtschaftlichkeitsanalyse im ILM.⁹⁸ Funktionen für die Klassifizierung und Verlagerung unstrukturierter Daten werden prototypisch implementiert.⁹⁹ Weiterhin werden Kostenfunktionen erstellt, die die relevanten Kosten, die bei der automatisierten Verwaltung der Daten entstehen, berechnen.¹⁰⁰ Die Implementierung erfolgt in Microsoft Excel.¹⁰¹ Der Schwerpunkt der prototypischen Implementierung liegt in der beispielhaften Umsetzung der Klassifizierung der In-

⁹² Vgl. Kapitel 5.1.3.2

⁹³ Vgl. Kapitel 5.2.1. Diese bilden gleichzeitig die Anforderungen für den in dieser Arbeit implementierten Prototyp und die durchgeführte Wirtschaftlichkeitsanalyse. Vgl. hierzu Kapitel 5.2 und Kapitel 5.3.

⁹⁴ In der Wirtschaftsinformatik stellt die Konstruktion von Konzepten und Modellen eine systematische und erschaffende Tätigkeit dar. Vgl. hierzu Fettke, Loos / Referenzmodelle / 27; Schütte / Modelle / 49. Der Aufbau von Strukturen bzw. Modellen erfolgt durch die Verwendung von Komponenten oder Bausteinen. Vgl. hierzu Jablonski, Böhm, Schulze / Anwendungen und Methoden / 65.

⁹⁵ Das Vorgehensmodell enthält die Teilphasen „Analyse der IS-Architektur“, „Beschreibung relevanter Datenbestände“, „Klassifizierung der Daten“ und „Verlagerung der Daten“. Vgl. hierzu Kapitel 5.1.2.

⁹⁶ Vgl. hierzu Kapitel 5.1.1. Die Untersuchung existierender ILM-Vorgehensmodelle ist Bestandteil der Literaturanalyse zum ILM.

⁹⁷ Vgl. hierzu Kapitel 1.1, Kapitel 1.2 und Kapitel 5.1.2.3. Das ILM-Vorgehensmodell wurde auf Basis der Literaturrecherche konstruiert. Vgl. hierzu Kapitel 5.1.2.

⁹⁸ Vgl. Kapitel 5.2.2, Abbildung 5-8

⁹⁹ Vgl. hierzu Peffers et al. / Research Methodology / 55; Kapitel 5.2.3.1.1; Kapitel 5.2.3.2.

¹⁰⁰ Vgl. Kapitel 5.2.3.3

¹⁰¹ Vgl. Anhang 4

formationen und der Analyse der Wirtschaftlichkeit des ILM. Es wird ein Simulationsmodell implementiert, um das Kosten- und Nutzenverhältnis für verschiedene Lebenszyklen, Datenmengen und Speicherebenen zu simulieren bzw. zu analysieren. Der Wert einer Information resultiert aus der Quantifizierung des Nutzungsgrads der Informationen bzw. der Zugriffe auf die Informationen im Zeitverlauf.¹⁰² Im Rahmen der Klassifizierung erfolgt die Berechnung der Klassen bzw. Klassenwerte, denen die Informationen entsprechend ihres Nutzungsgrads zugeordnet werden. Die Klassifizierungskriterien bzw. Lebenszyklen der Informationen, wie beispielsweise die Anzahl der Lese- und Schreibzugriffe auf unstrukturierte Daten, werden in dieser Arbeit nicht aus einer einzelnen Organisation oder Unternehmung erhoben. Vielmehr findet der Rückgriff auf Forschungsergebnisse¹⁰³ zu Verhaltensstudien zur Nutzung von Informationen statt.¹⁰⁴ Die in diesen Studien untersuchten Informationen sind beispielsweise unstrukturierte Daten aus Forschung und Entwicklung, E-Mails, Finanz- und Medizindaten aus verschiedenen Forschungseinrichtungen und Unternehmen unterschiedlicher Branchen.¹⁰⁵ Die betrachteten Lebenszyklen werden durch den Einsatz von stochastischem Rauschen und unterschiedlich starken Rauschgraden verändert.¹⁰⁶

- **Demonstration:** Die Demonstration ist dazu geeignet, die Funktionsweise der implementierten Funktionen der Wirtschaftlichkeitsanalyse nachzuweisen. Es wird untersucht, ob und inwieweit das konstruierte und implementierte Verfah-

¹⁰² Vgl. Kapitel 2.4.1, Kapitel 5.2.2.1

¹⁰³ Damit werden Aussagen zur Nutzung von Informationen erarbeitet, die unabhängig von bestimmten organisatorischen oder technischen Gegebenheiten einer einzelnen Unternehmung sind. Vgl. hierzu Frank / Formale Sprachen / 142-143; Satyanarayanan / Study Lifetimes / 99; Turczyk / ILM / 87; Venable / Design Science Research / 117; Venable / Framework for Design Science Research Activities / 186.

¹⁰⁴ Vgl. Kapitel 5.2.2.1.3

¹⁰⁵ Es werden sechs typische Lebenszyklen von Daten herausgearbeitet, die am häufigsten in den Forschungsprojekten identifiziert wurden. Der am häufigsten vorkommende Lebenszyklus beschreibt einen rapiden Rückgang der Zugriffshäufigkeiten, kurz nach der Entstehung der Daten. Altern die Daten im Zeitverlauf, so wird mit dem Voranschreiten ihres Lebenszyklus die Relevanz für den Geschäftsablauf entsprechend geringer. Die weiteren Lebenszyklen beschreiben jeweils steigende, konstante, saisonale und trend-saisonale Lebenszyklen bzw. Nutzungsgrade. Vgl. hierzu Kapitel 5.2.2.1.3.1. Auf Basis dieser Lebenszyklen werden Daten in Microsoft Excel generiert, welche die Nutzungsgrade für die Wirtschaftlichkeitsanalyse repräsentieren. Vgl. hierzu Kapitel 5.2.2.1.3.1 und Anhang 2.

¹⁰⁶ Vgl. Kapitel 5.2.2.1.3; Box, Muller / Random Normal Deviates / 610-611

ren zur Erreichung der Zielstellung der Arbeit beiträgt.¹⁰⁷ In der Demonstrationsphase werden durch den Einsatz der Simulationstechnik¹⁰⁸ die Kosten- und Nutzenbeziehungen der automatisierten Verwaltung der Daten analysiert.¹⁰⁹ Die Anschaffungs- und Betriebskosten der automatisierten Verwaltung unstrukturierter Daten werden mit Hilfe einer Kostenartenanalyse und den zuvor implementierten Funktionen ermittelt.¹¹⁰ Mit Hilfe der Methode der Kostenvergleichsrechnung¹¹¹ werden die Kosten der zu vergleichenden Alternativen¹¹² für die unterschiedlichen Lebenszyklen gegenübergestellt.¹¹³ Es wird für verschiedene Lebenszyklen von Informationen im Rahmen der Wirtschaftlichkeitsanalyse untersucht, unter welchen Bedingungen die Verwaltung unstrukturierter Daten langfristig die geringsten Kosten verursacht und in Bezug auf die Wirtschaftlichkeit¹¹⁴ den größten Nutzenbeitrag leistet.¹¹⁵ Es wird untersucht, inwieweit Variationen der Nutzungsgrade, dargestellt durch unterschiedlich starke Rauschgrade in den Lebenszyklen, Auswirkungen auf die Klassifizierungsergebnisse und das daraus resultierende Kosten- und Nutzenverhältnis haben.

- Evaluation: Ziel der Evaluation ist es, die Forschungsergebnisse hinsichtlich Stringenz und Nutzbarkeit bzw. praktischer Tauglichkeit zu untersuchen.¹¹⁶ In der Evaluationsphase wird untersucht, inwieweit die Ergebnisse der Demonstra-

¹⁰⁷ Vgl. hierzu Peffers et al. / Research Methodology / 55; Walls, Widmeyer, El Sawy / Design Theory / 36-59

¹⁰⁸ Vgl. Kolonko / Stochastische Simulation / 1-2; Küll, Stähly / Simulationsexperimente / 1-4

¹⁰⁹ Vgl. Kapitel 5.3

¹¹⁰ Hierfür werden marktübliche Preise zugrunde gelegt. Vgl. hierzu Anhang 5.3; Armbrust et al. / Cloud Computing / 2-25; Schadler / Email in the Cloud / 8-11; Suter / Marktübersicht Cloud / 36-37; Gantz, Reinsel / Extracting Value / 4; Reichman, Whiteley, Chi / File Storage Costs / 4-11; Komorowski / Cost per Gigabyte / ; Janukowicz / Solid State Storage / 2-4.

¹¹¹ Vgl. Kapitel 4.2.1.1

¹¹² Vgl. Kapitel 5.2.3.1.1.1 und Kapitel 5.3

¹¹³ Vgl. hierzu Kapitel 5.3 und Schmidt / Wirtschaftlichkeit / 49; Mensch / Investition / 44-45. Durch verschiedene Simulationsläufe bzw. Experimente werden Aussagen und Erkenntnisse erarbeitet, die auf die Wirklichkeit übertragbar sind. Vgl. hierzu VDI / Simulation / 2.

¹¹⁴ Mit Hilfe der Wirtschaftlichkeitsanalyse wird das nachhaltig günstigste Verhältnis zwischen Gewinn und Kapitaleinsatz sowie zwischen Kosten und Nutzen ermittelt (Vgl. hierzu Kapitel 4.1; Nebel / Produktionswirtschaft / 20; Heinrich, Stelzer / Informationsmanagement / 368-369).

¹¹⁵ Vgl. Kapitel 5.3

¹¹⁶ Vgl. Venable, Pries-Heje, Baskerville / Framework for Evaluation / 427

tionsphase zum Erreichen der Zielstellung¹¹⁷ der Arbeit beitragen. Es wird überprüft, ob die definierten Anforderungen an die Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM¹¹⁸ erfüllt werden. Mit Hilfe eines Merkmalsvergleichs werden die definierten Anforderungen mit den Ergebnissen der Demonstration abgeglichen.¹¹⁹ Hierbei handelt es sich um eine nachträgliche bzw. „ex post Evaluation“, bei der die Ergebnisse der Simulation ausgewertet werden.¹²⁰ Weiterhin wird erörtert, ob die erzielten Ergebnisse¹²¹ allein aus der durchgeführten Wirtschaftlichkeitsanalyse¹²² resultieren¹²³ und ob diese auf die Realität übertragbar sind, um die prototypische Implementierung bspw. in einem Unternehmen einzusetzen.

- **Kommunikation:** Im Anschluss an die Evaluationsphase wird die Kommunikation der Forschungsergebnisse empfohlen, um diese mit Hilfe von bspw. Veröffentlichungen der Forschungsgemeinschaft und Öffentlichkeit zugänglich zu machen.¹²⁴ Forschungsergebnisse zur vorliegenden Arbeit sind bereits zur Thematik automatisierte Verwaltung von Daten im ILM und Cloud Computing publiziert worden.¹²⁵ Teile der Ergebnisse dieser Arbeiten werden in Publikationen weiterer Autoren diskutiert.¹²⁶ Als ein Teil der Kommunikationsphase wird ebenfalls die Veröffentlichung der vorliegenden Arbeit verstanden.

¹¹⁷ Vgl. Kapitel 1.3

¹¹⁸ Vgl. Kapitel 5.2.1

¹¹⁹ Vgl. Peffers et al. / Research Methodology / 56

¹²⁰ Vgl. hierzu Venable, Pries-Heje, Baskerville / Framework for Evaluation / 432-433. Vgl. Kapitel 5.4 zur Diskussion weiterer Evaluationsmethoden.

¹²¹ Vgl. Kapitel 5.3

¹²² Die Wirtschaftlichkeitsanalyse basiert auf der prototypischen Implementierung. Vgl. hierzu Kapitel 5.2.3.

¹²³ Oder ob diese durch weitere Faktoren beeinflusst werden.

¹²⁴ Vgl. Hevner et al. / Design Science in IS / 75-105; Peffers et al. / Research Methodology / 54.

¹²⁵ Vgl. Matthesius, Stelzer / Analyse und Vergleich Konzepte im ILM / 471-482; Kosler, Matthesius, Stelzer / Klassifizierung von Informationen / 129-145; Cheptsov et al. / Scalable Cloud Platform / 8-12

¹²⁶ Vgl. hierzu bspw. Lehner / Wissensmanagement / 9; Turczyk / ILM / 12, 28; Heinrich, Stelzer / Informationsmanagement / 278-279; Wijnhoven, Amrit / Use-based file retention method / 4.

In Tabelle 1-1 sind zusammenfassend die Forschungsmethoden dargestellt, die innerhalb der einzelnen Phasen des Forschungsprozesses und der jeweiligen Kapitel dieser Arbeit verwendet werden.

Phase Forschungsprozess	Kapitel	Forschungsmethoden
Problemidentifikation und Motivation	Kapitel 1.1	Literaturanalyse
	Kapitel 2 bis Kapitel 4	Literaturanalyse
Zielfestlegung und Ableitung von Anforderungen	Kapitel 1.2	Literaturanalyse
	Kapitel 2.4	Literaturanalyse
	Kapitel 3.4	Literaturanalyse
	Kapitel 5.1.3.2	Literaturanalyse
	Kapitel 5.2.1	Literaturanalyse
Design und Entwicklung	Kapitel 5.1	Literaturanalyse, Konstruktion Vorgehensmodell
	Kapitel 5.2	Literaturanalyse Konstruktion und Implementierung Verfahren zur Wirtschaftlichkeitsanalyse Entwicklung Kostenmodell Erhebung Messwerte, stochastisches Rauschen
Demonstration	Kapitel 5.3	Simulation Kostenvergleichsrechnung
Evaluation	Kapitel 5.4	Merkmalsvergleich
Kommunikation	Kapitel 1.3.1	Publikationen
		Veröffentlichung der Arbeit

Tab. 1-1: Forschungsprozess und Forschungsmethoden

1.3.2 Methodische Vorgehensweise bei der Literaturanalyse

Ziel der Literaturanalyse ist es, den aktuellen Stand der Forschung zur Wirtschaftlichkeitsanalyse der automatisierten Klassifizierung unstrukturierter Daten im ILM zu ermitteln.¹²⁷ Den Schwerpunkt der Untersuchungen bildet die Analyse von Beiträgen zu den Themengebieten ILM¹²⁸, Vorgehensmodelle und Konzepte zur automatisierten Klassifizierung im ILM¹²⁹ sowie Wirtschaftlichkeitsanalyse im ILM¹³⁰. Begriff und Dimensionen des Cloud Computings¹³¹ sowie technische¹³² und rechtliche Anforderungen¹³³, die beim Einsatz von ILM und Cloud Computing zu beachten sind, sind ebenfalls Bestandteil der Literaturanalyse.

In Anlehnung an die Vorgehensweise zur Literaturanalyse nach Webster und Watson¹³⁴ werden relevante Beiträge aus Fach- und Journalartikeln, Sammelwerken, Monografien, Zeitschriften und Konferenzen identifiziert und analysiert. Die methodische Vorgehensweise zur Literaturanalyse besteht aus den folgenden Schritten:

- Auswahl von Zeitschriften und Konferenzen: Anhand von Rankings werden die wichtigsten bzw. relevanten deutsch- und englischsprachigen Zeitschriften und Konferenzen identifiziert.¹³⁵ Neben wissenschaftlichen Beiträgen werden zusätzlich Beiträge von Analysten und Speicherherstellern untersucht, da zum Konzept des ILM bislang wenig wissenschaftliche Literatur vorliegt.¹³⁶ Die Literatur zu weiteren Forschungsgebieten, die mit ILM in Beziehung stehen, wie bspw. zur

¹²⁷ Vgl. hierzu Fettke / State of the art / 257-260; Webster, Watson / Literature Review / 16. Vgl. hierzu Kapitel 1.2 und Kapitel 1.3.1 zur Zielsetzung der Arbeit und der Forschungsmethodik.

¹²⁸ Vgl. hierzu Kapitel 2, Begriff des Lebenszyklus und des ILM.

¹²⁹ Konzepte zur automatisierten Klassifizierung im ILM werden gegenübergestellt und miteinander verglichen. Die Gegenüberstellung der Konzepte erfolgt in Matrixform. Vgl. hierzu Webster, Watson / Literature Review / 17-18. Vgl. Kapitel 5.1.3.1.2. Es werden Anforderungen herausgearbeitet, die die Autoren an ihre jeweiligen Konzepte stellen. Hieraus werden Anforderungen an die automatisierte Klassifizierung unstrukturierter Daten im ILM abgeleitet. Vgl. hierzu Kapitel 5.1.3.2.

¹³⁰ Vgl. hierzu Kapitel 4. Begriff und Methoden zur Wirtschaftlichkeitsanalyse werden erörtert und es wird ein Bezug zum ILM hergestellt. Vgl. hierzu Kapitel 4.2.

¹³¹ Vgl. Kapitel 3

¹³² Vgl. hierzu Kapitel 2.4.2 und 3.4.1.

¹³³ Vgl. Kapitel 2.4.3 und 3.4.2

¹³⁴ Vgl. Webster, Watson / Literature Review / 15-16

¹³⁵ Vgl. Tabelle 1-3; Heinzl, Schoder, Frank / WI-Journallisten / 159-160; AIS / MIS Journal Ranking /

¹³⁶ Vgl. Kaiser, Smolnik, Riempp / ILM-Framework / 483-484; Turczyk / ILM / 11-14; Turczyk / Wertzuweisung / 459-460

Verwaltung von Speichermedien in Unternehmen bzw. zum Storage-Management, wird in die Analysen einbezogen.¹³⁷ Zur Analyse der rechtlichen Anforderungen werden die relevanten rechtlichen Vorschriften untersucht.¹³⁸ Für die Erörterung von bspw. Begriffsdefinitionen, Methoden und Dimensionen der Wirtschaftlichkeitsanalyse bzw. des Cloud Computings erfolgt zusätzlich die Analyse von Standardliteratur. Hierzu zählen Lehrbücher, Handbücher oder Nachschlagewerke.

- Durchsuchen der Zeitschriften und Konferenzen: Durch Verwendung von Schlagworten werden zentrale Datenbanken und Datenbankarchive der Zeitschriften und Konferenzen durchsucht. Bei der Suche in zentralen Datenbanken wird übergreifend über alle in der Datenbank vorhandenen Zeitschriften und Konferenzen gesucht. Die Suche der Schlagworte erfolgt in den Titeln, Zusammenfassungen und Schlüsselwörtern der Beiträge. Zusätzlich erfolgt die Suche mit Hilfe der Suchmaschinen-Dienste „Google“, „Google Scholar“ und „Microsoft Academic Search“. Auf diese Weise werden auch Beiträge von Analysten, Beratungsgesellschaften und Speicherherstellern identifiziert. In Tabelle 1-2 sind die verwendeten Schlagworte abgebildet.

Schlagworte (englisch)	Schlagworte (deutsch)
Information Lifecycle Management	Informationslebenszyklusmanagement
Data Lifecycle Management	Datenlebenszyklusmanagement
Information Lifecycle	Informationslebenszyklus
Data Lifecycle	Datenlebenszyklus

Tab. 1-2: Schlagworte Literaturanalyse

¹³⁷ Vgl. hierzu Webster, Watson / Literature Review / 15-16. Bspw. identifiziert durch Vor- und Rückwärtssuche in den Beiträgen zum ILM.

¹³⁸ Vgl. Kapitel 2.4.3 und Kapitel 3.4.2

Gesucht wird nach den Begriffen „Information Lifecycle Management“ und „Data Lifecycle Management“ in englischer und deutscher Sprache.¹³⁹ Bei den englischen Schlagworten wird durch das Voranstellen von „Information“ und „Data“ eine Abgrenzung zu bspw. den Themengebieten „Product-Lifecycle“ und „Software-Lifecycle“ vorgenommen. Da in den Titeln und Zusammenfassungen der durchsuchten Beiträge oft der Begriff „Management“ von den Autoren weggelassen wird, erfolgt zusätzlich die Suche anhand der Schlagworte „Information Lifecycle“ und „Data Lifecycle“. Die Vorgehensweise bei den deutschen Schlagworten erfolgt analog.

- Analyse und Selektion von Beiträgen: Die bei der Suche gefundene Literatur wird untersucht und relevante Beiträge werden aus der gesamten Treffermenge selektiert und analysiert. Hierzu werden die Volltexte der Beiträge untersucht. Die Artikel, die einen Beitrag zum Forschungsgegenstand dieser Arbeit leisten, werden in die weiteren Betrachtungen einbezogen. Von Interesse ist, ob in den Beiträgen neben Begriffsdefinitionen und Vorgehensmodellen zum ILM zusätzlich Konzepte zur Klassifizierung unstrukturierter Daten beschrieben werden oder ob in der Literatur zum ILM ein Bezug zur Wirtschaftlichkeitsanalyse oder zum Cloud Computing hergestellt wird.¹⁴⁰ Beiträge, die vorwiegend Konzepte oder Ergebnisse bereits berücksichtigter Arbeiten beschreiben, werden nicht berücksichtigt. Gleiches gilt für Beiträge, die vorwiegend eingesetzte Datenbankmanagementsysteme oder Datenanalysewerkzeuge erörtern.¹⁴¹ Zusätzlich werden durch Vor- und Rückwärtssuche in den berücksichtigten Beiträgen referenzierte oder inhaltlich verwandte Artikel identifiziert.¹⁴² Diese Artikel werden wiederum analysiert und gefiltert.¹⁴³ Die durchsuchten Datenbanken, Zeitschriften und Konferenzen sind in Tabelle 1-3 enthalten.¹⁴⁴

¹³⁹ In der Literatur werden beide Begriffe häufig gleichgesetzt. Vgl. Kapitel 1.1.

¹⁴⁰ Vgl. Tabelle 1-4

¹⁴¹ Der Fokus dieser Arbeit liegt auf der Klassifizierung von Daten im ILM und der damit verbundenen Wirtschaftlichkeit.

¹⁴² Webster, Watson / Literature Review / 16

¹⁴³ Bei der Identifikation weiterer Zeitschriften durch die Vor- und Rückwärtssuche werden auch die betreffenden Archive analysiert.

¹⁴⁴ Abgebildet sind Zeitschriften und Konferenzen, die im Hinblick auf ILM untersucht wurden.

Datenbanken / Archive	Publikationen¹⁴⁵	Zeiträume¹⁴⁶	Analy-sierte Beiträge	Berück-sichtigte Beiträge
ACM Digital Library	<u>Zeitschriften:</u> Journal of Advanced Engineering Informatics; Journal of Computers & Security; Semantic Web Journal; IEEE Multimedia; IEEE Transactions on Services Computing; International Journal of Business Information Systems; International Journal of Information Management; International Journal on Digital Libraries; Transactions on Computer Research	2002-2014 1982-2014 2010-2014 1994-2014 2008-2014 2005-2014 1986-2014 1997-2014 1999-2014	10	3
	<u>Konferenzen:</u> Proceedings ACM; SIGMOD; IEEE Computer Society; Computer Society India; WSEAS	Bis 2014	42	9
Archive	<u>Zeitschriften:</u> Information Systems Research; European Journal of Information Systems; HMD Praxis der Wirtschaftsinformatik; Wirtschaftsinformatik	1990-2014 1991-2014 2007-2014 1995-2014	6	2
Directory of Open Access Journals (DOAJ) ¹⁴⁷	<u>Zeitschriften:</u> International Journal of Digital Curation; Journal of Open Research Software	2008-2014 2013-2014	5	2

¹⁴⁵ Es werden die Publikationen aufgeführt, für die die Schlagwortsuche Treffer ergab.

¹⁴⁶ Die Startzeitpunkte sind die jeweils ersten Erscheinungsjahre der Publikationen. Das Enddatum des Analysezeitraums ist der 31.12.2014.

¹⁴⁷ Beiträge zu Konferenzen sind im DOAJ nicht enthalten.

Datenbanken / Archive	Publikationen	Zeiträume	Analy-sierte Beiträ-ge	Berück-sichtigte Beiträge
Science Direct	<u>Zeitschriften:</u> Journal of China Universities of Posts and Telecommunications; Journal of Systems and Software; Decision Support Systems; Journal of Fusion Engineering and Design; Journal of Computer Standards and Interfaces	2006-2014 1980-2014 1985-2014 1987-2014 1986-2014	6	0
Vor- und Rückwärtssuche	<u>Zeitschriften:</u> Journal of Advances in Soft Computing & its Applications Data Engineering Bulletin; ACM/SIGSMALL PC Notes; Communications of the ACM; IBM Systems Journal; McKinsey Quarterly; IEEE Computer; CAD-CAM Report; ECMguide Magazin; Information, Wissenschaft und Praxis	2009-2014 1977-2014 1985-2014 1958-2014 2006-2014 1964-2014 1970-2014 1982-2014 2004-2014 2005-2014	33	10
	<u>Konferenzen:</u> Proceedings USENIX; International Conference Distributed Computing Systems; ACM SIGOPS; ACM SIGMOD; MKWI; Advances in Web-Age Information Management; DECUS IT Symposium; IEEE Symposium Mass Storage Systems; Autonomic Computing; ISSE; AKWI; SIGSVC; European Conference Research and Technology; Workshop Systems Dependability	Bis 2014	47	19

Tab. 1-3: Literaturanalyse zum ILM (Zeitschriften und Konferenzen)

Es werden Angaben zu Analysezeiträumen sowie der Anzahl analysierter und berücksichtigter Beiträge gemacht. Das Startdatum des Analysezeitraums wird nicht eingeschränkt, es werden rückwirkend die Jahrgänge durchsucht, die für

die jeweiligen Zeitschriften und Konferenzen existieren bzw. in den Datenbanken vorhanden sind. Die Literaturanalyse wurde im Jahr 2014 abgeschlossen, bis dahin erschienene Beiträge wurden untersucht. Aus insgesamt 149 analysierten Beiträgen wurden 45 Beiträge¹⁴⁸ berücksichtigt. Davon wurden 16 relevante Beiträge aus der Schlagwortsuche gefunden. Durch die Vor- und Rückwärtssuche und die zusätzliche Untersuchung von Beiträgen der Themengebiete Storage Management und Verwaltung von Unternehmensdaten¹⁴⁹ wurden weitere 29 Beiträge identifiziert. Die untersuchte Standardliteratur und die analysierten Beiträge aus Zeitschriften und Konferenzen zu bspw. Begriff, Dimensionen und Methoden des Cloud Computings¹⁵⁰ und der Wirtschaftlichkeitsanalyse werden nicht in Tabelle 1-3 aufgelistet. Diese werden in den entsprechenden Kapiteln der Arbeit in Quellenverweisen zitiert.¹⁵¹ Gleiches gilt für Beiträge, die mit Hilfe von Suchmaschinen-Diensten gefunden wurden.

- Dokumentation und Nutzung der Ergebnisse: In der Literatur identifizierte Konzepte, Methoden und Modelle werden dokumentiert, miteinander verglichen und analysiert.¹⁵² Offene Forschungsthemen werden identifiziert und relevante Problemstellungen werden abgeleitet.¹⁵³ Aufbauend auf den Erkenntnissen der Literaturanalyse werden Anforderungen an das in dieser Arbeit zu konstruierende Verfahren zur Wirtschaftlichkeitsanalyse im ILM abgeleitet.

Eine Übersicht zur dokumentierten und analysierten Literatur zum ILM aus den jeweiligen Zeitschriften und Konferenzen ist in Tabelle 1-4 enthalten.

¹⁴⁸ Vgl. Kapitel 1.3.2; Abbildung 1-4

¹⁴⁹ Da in den meisten Beiträgen zum ILM Literaturverweise zum Storage Management und zur Verwaltung von Unternehmensdaten enthalten waren, wurde die Literaturanalyse sowie die Vor- und Rückwärtssuche auch für relevante Beiträge aus diesen Themengebieten durchgeführt.

¹⁵⁰ Vgl. Kapitel 3

¹⁵¹ Vgl. Kapitel 2.4.3; Kapitel 3; Kapitel 4

¹⁵² Hierzu eignet sich die Verwendung einer Konzept-Matrix. Vgl. hierzu Webster, Watson / Literature Review / 17-18.

¹⁵³ Vgl. Kapitel 1.3.1

Autoren (Kurztitel)	Jahr	Automati- sierte Klassifi- zierung	Wirt- schaftlich- keitsanaly- se	Unstruk- turierte Daten	Vorge- hensmo- delle	Rechtliche Anforde- rungen	Cloud
Alatorre et al. (Intelli- gent ILM)	2014	X	-	X	-	-	-
Du et al. (Optimal Storage)	2014	-	-	X	-	-	X
Jeong et al. (Effective Storage)	2014	X	-	-	-	-	-
Lenhardt et al. (Data Lifecycle)	2014	-	-	X	-	-	-
Badam, Kumari, Srimi- vasan (ILM healthcare)	2013	X	-	-	-	-	-
Möller (Lifecycle models)	2013	-	-	X	-	-	-
Lee et al. (ILM with RFID)	2013	X	-	X	-	-	-
Steinbinder, Bedau, Löw (Datenbanktech- nologie für das ILM)	2011	X	-	-	-	X	-
Wijnhoven, Amrit (Use-based file reten- tion method)	2010	X	-	X	-	-	-
Liu, Wang, Quan (ILM Architecture)	2009	-	-	X	-	-	-
Al-Fedaghi (ILM)	2008	X	-	X	-	-	-
Doukas, Ntaikos, Bardis (ILM Applica- tions)	2008	-	-	X	-	X	-
Kaiser, Smolnik, Riempp (ILM- Framework)	2008	-	-	X	-	X	-
Kaiser, Smolnik, Riempp (Verbesserte Compliance ILM)	2008	-	-	X	-	X	-
Kaplan, Roy, Srimi- vasaraghavan (Demand data storage)	2008	-	-	X	-	-	-
Turczyk (Wertzuwei- sung)	2008	X	-	X	-	-	-
Agrawal et al. (Study File-System Metadata)	2007	-	-	X	-	-	-
Bernard (ILM Securi- ty)	2007	-	-	X	-	X	-
Chandra, Gehani, Yu (Automated Storage)	2007	X	-	X	-	-	-
Kanakamedala, Kaplan, Srinivasara- ghavan (Data storage)	2007	-	-	X	-	-	-
Mont, Beato (Obliga- tion Policies)	2007	-	-	X	-	X	-
Abd-El-Malek et al. (Self-storage)	2006	X	-	X	-	-	-
Lotlikar, Mohania (Policies ILM)	2006	X	-	X	-	X	-
Rüdiger (Stückwerk ILM)	2006	-	-	X	-	-	-

Autoren (Kurztitel)	Jahr	Automati- sierte Klassifi- zierung	Wirt- schaftlich- keitsanaly- se	Unstruk- turierte Daten	Vorge- hensmo- delle	Rechtliche Anforde- rungen	Cloud
Mont (Privacy-aware ILM)	2006	-	-	X	-	X	-
Abd-El-Malek et al. (Cluster-based storage)	2005	X	-	X	-	-	-
Bhagwan et al. (Management of Data Storage)	2005	X	-	X	-	-	-
Chen (Information Valuation)	2005	X	-	X	-	-	-
Mont et al. (Privacy Policy Enforcement)	2005	-	-	X	-	X	-
Turczyk (Organisation ILM)	2005	X	-	X	-	-	-
Verma et al. (Very Large File Systems)	2005	X	-	X	-	-	-
Douglis et al. (Delete-Optimized Storage)	2004	X	-	X	-	-	-
Liu, Li, Zhang (Dispatching Methods)	2004	-	-	X	-	-	-
Mesnier et al. (File classification)	2004	X	-	X	-	-	-
Reiner et al. (ILM EMC)	2004	X	-	X	X	-	-
Zadok et al. (Reducing Storage Management Costs)	2004	X	-	X	-	-	-
Gießelbach (ILM)	2003	X	-	X	X	X	-
Witten et al. (Documents and Metadata)	2002	-	-	X	-	-	-
Hillyer, Silberschatz (Tertiary storage systems)	1996	-	-	X	-	-	-
Miller, Katz (Analysis File Migration)	1993	-	-	X	-	-	-
Bennett, Bauer, Kinchlea (NFS environments)	1992	-	-	X	-	-	-
Gavish, Sheng (Dynamic file migration)	1990	X	-	X	-	-	-
Merrill, Thanhardt (Mass Storage)	1990	-	-	X	-	-	-
Gelb (System-Managed Storage)	1989	X	-	X	-	-	-
Lawrie, Randal; Barton (File Migration)	1982	X	-	X	-	-	-

Tab. 1-4: Dokumentation Literaturanalyse ILM (Zeitschriften und Konferenzen)

In den Beiträgen zum ILM wird untersucht, inwiefern die Themengebiete automatisierte Klassifizierung, Wirtschaftlichkeitsanalyse, Verwaltung unstrukturierter Daten, rechtliche Anforderungen und Cloud Computing von den Autoren bearbeitet wurden. Weiter-

hin wird untersucht, ob Vorgehensmodelle zum ILM publiziert wurden. Somit wird der Forschungsstand zum ILM ermittelt. In 24 der insgesamt 45 untersuchten Beiträge wird die automatisierte Klassifizierung von Daten diskutiert. Einige Autoren weisen lediglich auf die Notwendigkeit zur automatisierten Klassifizierung im ILM hin, andere entwickeln und evaluieren Konzepte zur Klassifizierung der Informationen.¹⁵⁴ In zwei Beiträgen werden Vorgehensmodelle zum ILM vorgestellt.¹⁵⁵ Diese beziehen sich hauptsächlich auf die Verwaltung von Speichermedien während der einzelnen Phasen des Informationslebenszyklus. Anhand welcher Kriterien Informationen bewertet bzw. klassifiziert werden können, wird in den Modellen nicht beschrieben. Um eine Wirtschaftlichkeitsanalyse im ILM durchführen zu können, muss das Vorgehen bei der Umsetzung von ILM und bei der Klassifizierung der Informationen beschrieben werden.¹⁵⁶ Aufbauend auf den in der Literatur beschriebenen Vorgehensmodellen wird hierzu ein ILM-Vorgehensmodell in dieser Arbeit konstruiert.¹⁵⁷

In keinem der im Rahmen der Literaturanalyse berücksichtigten Beiträge wird ein Verfahren zur Analyse der Wirtschaftlichkeit des ILM beschrieben, entwickelt oder angewendet.¹⁵⁸ Es erfolgt lediglich der Hinweis, dass ILM zu einer Senkung von Speicher- und Administrationskosten beitragen soll. Wie und in welchem Maß dies erfolgt, wird von den Autoren nicht untersucht. Diese Forschungslücke wird durch die vorliegende Arbeit dahingehend geschlossen, dass erstmals Aussagen zur Wirtschaftlichkeit von ILM getroffen werden und ein Verfahren zur Wirtschaftlichkeitsanalyse zur Verwaltung unstrukturierter Daten entwickelt und prototypisch implementiert wird.¹⁵⁹ Es wird aufgezeigt, unter welchen Rahmenbedingungen ILM wirtschaftlich ist. Lediglich drei Beiträge beschäftigen sich mit der Verwaltung strukturierter Daten.¹⁶⁰ Die Verwaltung unstrukturierter Daten im ILM steht im Vordergrund.¹⁶¹ In zehn Beiträgen wird die Bedeu-

¹⁵⁴ Vgl. hierzu Kapitel 5.1.3 zur Diskussion der Konzepte zur automatisierten Klassifizierung.

¹⁵⁵ Weitere zwei Beiträge, in denen Vorgehensmodelle publiziert werden, wurden durch die Suche mit Hilfe von Suchmaschinen-Diensten identifiziert. Vgl. hierzu Kapitel 5.1.1. Diese Beiträge sind nicht in Tabelle 1-4 enthalten.

¹⁵⁶ Bspw. zur Abschätzung von Aufwand und Nutzen bei der Informationsklassifizierung.

¹⁵⁷ Vgl. Kapitel 5.1.2

¹⁵⁸ Vgl. Tabelle 1-4

¹⁵⁹ Vgl. Kapitel 5.2

¹⁶⁰ Vgl. Kapitel 4.3

¹⁶¹ Vgl. Kapitel 1.1

tung von rechtlichen Anforderungen diskutiert, die bei der Verwaltung und Verlagerung von Informationen berücksichtigt werden müssen. Es wird beschrieben, dass rechtliche Vorgaben von den Anwendern und Administratoren in Form von Regeln definiert werden müssen. Diese Regeln müssen bei der Klassifizierung der Informationen berücksichtigt werden. Die automatisierte Erstellung und Aktualisierung von Regeln zu rechtlichen Anforderungen, bspw. auf der Grundlage von Gesetzesänderungen, ist nicht möglich.¹⁶² Eine Betrachtung, welche rechtlichen Vorschriften beim Einsatz von ILM relevant sind, wird in der Forschung bislang vernachlässigt. In dieser Arbeit werden die relevanten rechtlichen Vorschriften zur Verwaltung von Informationen im ILM, unter Berücksichtigung des Cloud Computings, erörtert. Die Nutzung des Cloud Computings im Kontext von ILM wird lediglich in einem der analysierten Beiträge diskutiert.¹⁶³ Inwieweit die Nutzung von unternehmensinternen Speicherebenen und Cloud-Diensten zu einer wirtschaftlich sinnvollen Verwaltung der Daten im ILM führen kann, wird in diesem Beitrag nicht beschrieben.

Im Zeitraum 2004-2008 sind Konzepte entstanden, um unstrukturierte Daten automatisiert zu klassifizieren.¹⁶⁴ Teilweise jedoch noch nicht unter dem Schlagwort ILM.¹⁶⁵ Ein weiteres Konzept stammt aus dem Jahr 2014.¹⁶⁶ In den Jahren 2005 und 2008 wurde der Begriff des ILM geprägt. In diesem Zeitraum sind die meisten Publikationen entstanden. Von 2009 bis 2012 sind kaum Publikationen zum ILM entstanden. Seit 2013 und 2014 werden wieder mehr Publikationen veröffentlicht, ILM gewinnt wieder zunehmend an Bedeutung. Hier lässt sich ein Zusammenhang mit dem ebenfalls in diesen Jahren aufkommenden Begriff „Big Data“ vermuten. Die Problematik wachsender Da-

¹⁶² Dies deckt sich mit den Erkenntnissen aus Publikationen der Jahre 2013 und 2014 zur Thematik Sicherheit in Smart Grids. Vgl. hierzu Park / Privacy Policy Grid / 91-100; Park, Kim / Policy energy management / 653-664.

¹⁶³ Im betreffenden Beitrag wird erörtert, wie Empfehlungssysteme von Videoinhalten einer Cloud-Plattform gestaltet sein sollten. Es wird die Verwaltung der Daten in unterschiedlichen Speicherebenen empfohlen. Vgl. hierzu Du et al. / Optimal Storage / 730-740.

¹⁶⁴ Nicht alle Beiträge, in denen Konzepte zur automatisierten Klassifizierung diskutiert werden, sind in Tabelle 1-4 enthalten. Das bedeutet, dass die nicht enthaltenen Beiträge nicht in den Datenbankarchiven von Zeitschriften und Konferenzen gefunden wurden, sondern über die Suche mit Hilfe von Suchmaschinen. Vgl. hierzu die Methodik und Dokumentation zur Literaturanalyse Kapitel 1.3.2. Eine Übersicht zu den in der Literaturanalyse identifizierten Konzepten befindet sich in Kapitel 5.1.3.

¹⁶⁵ Von den Autoren wurden Begriffe verwendet wie „self storage“ oder „storage automation“.

¹⁶⁶ Vgl. Kapitel 5.1.3

tenmengen sowie steigender Speicher- und Administrationskosten rückt stärker in den Vordergrund.¹⁶⁷

1.4 Aufbau der Arbeit

Die vorliegende Arbeit liefert eine Literaturanalyse zu den Konzepten ILM und Cloud Computing sowie einen wissenschaftlichen Beitrag dazu, sowohl rechtliche als auch funktionale und technische Anforderungen an ILM und Cloud Computing zu erörtern.¹⁶⁸ Zu Beginn der Arbeit wird auf die Begriffe ILM, unstrukturierte Daten und Cloud Computing eingegangen. Die Begriffe ILM und Cloud Computing, die jeweiligen technischen, funktionalen und rechtlichen Anforderungen werden in den Kapiteln 2 und 3 erörtert. Die Methoden der Wirtschaftlichkeitsanalyse werden in Kapitel 4 vorgestellt. Es wird ein Bezug zwischen der Wirtschaftlichkeitsanalyse und ILM hergestellt. In Kapitel 5 erfolgt die Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM. Hierzu werden zunächst publizierte ILM-Vorgehensmodelle vorgestellt. Wesentliche Teilaufgaben des ILM werden anhand des konstruierten ILM-Vorgehensmodells erörtert. Das Cloud Computing wird als eine zusätzliche Möglichkeit zur Verwaltung unstrukturierter Daten in das ILM-Vorgehensmodell integriert. Anschließend wird das Verfahren für die Wirtschaftlichkeitsanalyse zur Verwaltung unstrukturierter Daten im ILM konstruiert, demonstriert und evaluiert. Beendet wird die Arbeit durch eine Zusammenfassung sowie eine kritische Würdigung der Ergebnisse. Ein Ausblick gibt Hinweise und Verbesserungsvorschläge hinsichtlich der Wirtschaftlichkeitsanalyse unstrukturierter Daten im Information Lifecycle Management.

¹⁶⁷ Vgl. McKinsey / Challenge Big Data / 2-3; Krishnan / Big Data / 3-5

¹⁶⁸ Vgl. Kapitel 2, Kapitel 3

2 Information Lifecycle Management (ILM)

2.1 Begriff des Lebenszyklus

Der Zyklus, abgeleitet aus dem griechischen „Kyklos“, beschreibt ein „periodisch ablaufendes Geschehen“¹⁶⁹. Ein Lebenszyklus beschreibt den Gang der Entwicklung eines Objektes über seine gesamte Lebensdauer hinweg.¹⁷⁰ Mit der Planung, Überwachung und Steuerung von Lebenszyklen befasst sich das Lebenszyklusmanagement.¹⁷¹ Die Lebenszyklustheorie, die zunächst als Instrument zur Analyse des internationalen Handels¹⁷² und später zur Analyse der dynamischen Entwicklung von Produkten im Markt¹⁷³ diente, bildet die Grundlage für das Lebenszyklusmanagement. Lebenszyklusmodelle gliedern sich typischerweise in verschiedene Phasen.¹⁷⁴ Der Produktlebenszyklus nimmt beispielsweise die Einordnung eines Produkts in die Phasen Einführungs-, Wachstums-, Reife- und Sättigungsphase vor.¹⁷⁵ Anhand dieser Phasen kann eine lebenszyklusbezogene Produkt- und Programmplanung erfolgen. Ein weiteres Beispiel ist der Softwarelebenszyklus. Dieser beschreibt die Phasen der Entwicklung, Implementierung, Nutzung, Wartung und Stilllegung der Software.¹⁷⁶ Im Rahmen dieser Arbeit ist der Lebenszyklus von Informationen relevant, welcher ebenfalls in bestimmte Phasen gegliedert ist.¹⁷⁷ Wie in Abbildung 2-1 dargestellt, besteht dieser aus den Phasen Erstellung, Verarbeitung und Nutzung, Aufbewahrung sowie Löschung von Informationen.¹⁷⁸

¹⁶⁹ Vgl. Brockhaus Enzyklopädie / Zyklus / 676

¹⁷⁰ Vgl. Balzert / Software-Technik / 1098

¹⁷¹ Vgl. Heinrich, Stelzer / Informationsmanagement / 272

¹⁷² Vgl. Posner / International Trade / 323-341

¹⁷³ Vgl. Polli, Cook / Product Life Cycle / 385-400

¹⁷⁴ Vgl. Alisch / Wirtschaftslexikon / 1876

¹⁷⁵ Vgl. Wöhe, Döring / Betriebswirtschaftslehre / 140

¹⁷⁶ Vgl. Fink, Schneidereit, Voß / Wirtschaftsinformatik / 178-182

¹⁷⁷ Vgl. Möller / Lifecycle models / 67-88

¹⁷⁸ Vgl. Lenhardt et al. / Data Lifecycle / 2 ; Petrocelli / Information Lifecycle Management / 194

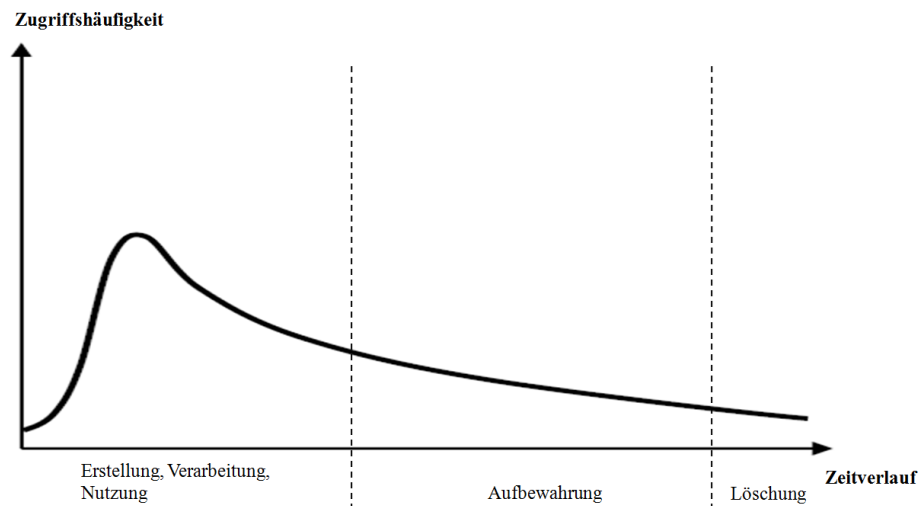


Abb. 2-1: Lebenszyklus von Informationen

Abbildung 2-1 zeigt ein Beispiel für einen Lebenszyklus von Informationen.¹⁷⁹ Bei der Erstellung, Verarbeitung und Nutzung der Informationen wird in der Regel häufig auf diese zugegriffen. Altern die Informationen, so wird mit dem Voranschreiten ihres Lebenszyklus die Relevanz für den betrieblichen Ablauf immer geringer. Aufgrund betrieblicher oder rechtlicher Vorgaben müssen die Informationen langfristig aufbewahrt werden.¹⁸⁰ Ist eine Aufbewahrung nicht mehr erforderlich, werden die Informationen gelöscht. Neben dem in Abbildung 2-1 dargestellten beispielhaften Lebenszyklus von Informationen sind weitere Lebenszyklen denkbar. Darauf wird in Kapitel 2.4.2 näher eingegangen.

2.2 Begriff des ILM

Seine Ursprünge hat das ILM im Datenmanagement. Technologische Vorgänger sind das System-Managed-Storage¹⁸¹, Hierarchical Storage Management (HSM)¹⁸² und Storage Resource Management (SRM)¹⁸³. Das Konzept des ILM stammt vorwiegend aus dem nicht-akademischen Umfeld. Es existieren zahlreiche Begriffserklärungen bzw. Definitionen in der Literatur sowie von Speicherherstellern, Beratungshäusern und Analysten. Petrocelli definiert ILM als ein Konzept, welches Informationen entsprechend

¹⁷⁹ Vgl. Chen / Information Valuation / 8; Inmon / ILM Data Warehousing / 5-6; Liu, Wang, Quan / ILM Architecture / 165-169; Kapitel 2.4.2

¹⁸⁰ Vgl. hierzu Kapitel 2.4.3

¹⁸¹ Vgl. Gelb / System-Managed Storage / 77-103

¹⁸² Vgl. Chen / Storage Hierarchies / 277-282

¹⁸³ Vgl. Clark / Storage Networks / 532-534

ihres Werts und ihrer Nutzung, auf dem jeweils kostengünstigsten Speichermedium bereitstellt.¹⁸⁴ Auch Organisationen wie die Storage Networking Industry Association (SNIA) und der Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V. (BITKOM) beschäftigen sich mit ILM. Die SNIA definiert ILM als einen Vorgang bzw. Prozess, welcher durch Einsatz geeigneter Hard- und Software den Wert einer Information ermittelt und entsprechend des Wertes möglichst kostengünstig verwaltet.¹⁸⁵ Der BITKOM versteht unter ILM „eine Regelmaschine [die] unter Berücksichtigung der Vorgaben aus den Geschäftsprozessen und der Bewertung der Kostenstrukturen der Speicherhierarchie in einem Optimierungsprozess den bestgeeigneten Speicherplatz für die verwalteten Informationsobjekte“¹⁸⁶ bestimmt. Nach Maier, Hädrich und Peinl liegt die Zielstellung von ILM darin, Informationen, wie Textdokumente, aufbereitete Grafiken, E-Mails oder Daten in einer Datenbank, die nicht für den täglichen Geschäftsablauf relevant sind, automatisch von den kostenintensiven Speichermedien der operativen Systeme zu entfernen und diese im Sinne einer effizienten Verwaltung auf kostengünstigere Speichermedien auszulagern oder zu löschen.¹⁸⁷ In einem Punkt sind diese Definitionen deckungsgleich: ILM steht insbesondere für die Klassifizierung und die (das) kostengünstige Verwaltung¹⁸⁸ (Management) von Informationen. Damit wird in der Literatur ILM definiert als die Verwaltung und Bereitstellung von Informationen, entsprechend der Unternehmensanforderungen und auf den jeweils sinnvollsten Speichermedien.¹⁸⁹ ILM wird in Bezug auf die Verwaltung von Informationen auf unternehmensinterne Speichermedien begrenzt. Die Nutzung des Cloud Computings, bspw. als Dienst über das Internet, kann jedoch kostengünstiger sein, als die Ver-

¹⁸⁴ Vgl. Petrocelli / Information Lifecycle Management / 180-181

¹⁸⁵ Vgl. SNIA / ILM Definition / 5

¹⁸⁶ Vgl. BITKOM / Leitfaden ILM / 4. Ein Geschäftsprozess ist eine zusammengehörige Abfolge von Unternehmensverrichtungen zum Zweck der Leistungserstellung (Vgl. hierzu Scheer / Geschäftsprozess / 3).

¹⁸⁷ Vgl. Maier, Hädrich, Peinl / Knowledge Infrastructures / 116

¹⁸⁸ Unter Verwaltung ist bspw. die Verlagerung oder das Kopieren der Informationen auf ein anderes Speichermedium oder die Löschung von Informationen zu verstehen. Im Rahmen dieser Arbeit wird der Begriff Management mit dem Begriff Verwaltung gleichgesetzt. Management im Sinne der Informationsverarbeitung sorgt für den „bestmöglichen Einsatz der Ressource Information“ (Vgl. hierzu Krcmar / Informationsmanagement / 1).

¹⁸⁹ Vgl. Abd-El-Malek et al. / Cluster-based storage / 59-72; Bhagwan et al. / Management of Data Storage / ; Glazer / Measuring Information / 99-110; Maier, Hädrich, Peinl / Knowledge Infrastructures / 109, 116; Mont / Privacy-aware ILM / 405-415; Petrocelli / Information Lifecycle Management / 180-181; Vgl. Kapitel 2.2

waltung der Informationen auf unternehmenseigenen Speichermedien.¹⁹⁰ Aus diesem Grund sollte das ILM erweitert werden, indem das Cloud Computing, als zusätzliche Möglichkeit zur Verwaltung von Informationen, mit einbezogen wird. In dieser Arbeit wird unter ILM die Verwaltung von Informationen in der jeweils sinnvollsten Speicherebene verstanden.¹⁹¹ Der Begriff „Speicherebene“ verdeutlicht, dass nicht nur Speichermedien innerhalb eines Unternehmens betrachtet werden. Vielmehr können Informationen auch außerhalb von Unternehmensgrenzen verwaltet werden, bspw. in der IT-Infrastruktur einer Public Cloud.¹⁹² ILM sorgt für eine Klassifizierung der Informationen, um zunächst die Relevanz für den Geschäftsablauf zu ermitteln. Es muss also der jeweilige Klassifizierungswert¹⁹³ der Informationen im Verlauf ihres Lebenszyklus zugewiesen werden. Des Weiteren müssen geeignete Speichertechnologien zur Verfügung gestellt werden, um eine kostengünstige Speicherung, entsprechend des klassifizierten Wertes der Informationen, zu ermöglichen. In Kapitel 2.3 werden die Eigenschaften von unstrukturierten Daten vorgestellt, anschließend werden in Kapitel 2.4 Anforderungen an ILM definiert.

2.3 Unstrukturierte Daten im Kontext von ILM

Bezüglich der Struktur von Daten können verschiedene Grade unterschieden werden. In der Literatur existieren in diesem Zusammenhang Begriffe wie strukturiert, semistrukturiert oder unstrukturiert.¹⁹⁴ Strukturierte Daten besitzen eine feste und vorgegebene Struktur. Einzelne Datenfelder haben eine definierte Länge und ein definiertes Format. Sie besitzen einheitliche Attribute mit definierten Wertebereichen. Zu den strukturierten Daten gehören die Daten aus operativen Datenbanken oder Data-Warehouse-Systemen¹⁹⁵. Im Gegensatz zu den strukturierten Daten existieren bei unstrukturierten

¹⁹⁰ Wenn Informationen im Cloud Computing als Dienst über das Internet verwaltet werden, können die Kosten für die Administration unternehmenseigener Speichermedien weiter gesenkt werden. Vgl. hierzu Schadler / Email in the Cloud / 3-5.

¹⁹¹ Hierzu muss die bislang in der Literatur beschriebene Speicherhierarchie für ILM (Vgl. Kapitel 2.4.2) erweitert werden, indem das Cloud Computing integriert wird. Dieser Sachverhalt wird in Kapitel 5.1.2 beschrieben. Es wird ein Vorgehensmodell für das ILM entwickelt und neben den klassischen Speicherebenen „Online“, „Nearline“ und „Offline“ zusätzlich das Cloud Computing betrachtet (Vgl. hierzu Abbildung 5-5).

¹⁹² Vgl. Kapitel 3.2.1

¹⁹³ Vgl. Kapitel 2.4.1

¹⁹⁴ Vgl. Abiteboul / Semi-Structured Data / 1-18; Stahlknecht, Hasenkamp / Wirtschaftsinformatik / 135-137

¹⁹⁵ Vgl. Inmon / Data Warehouse / 31

Daten keine vorgegebenen Anordnungen von Datenfeldern oder Wertebereiche für Attribute.¹⁹⁶ Zu den unstrukturierten Daten zählen beispielsweise E-Mails, Dokumente, Grafiken oder Videos.¹⁹⁷ Eine Zwischenform im Hinblick auf den Strukturierungsgrad von Daten sind semistrukturierte Daten.¹⁹⁸ Für diese Daten ist keine feste Struktur vorgeschrieben, jedoch tragen sie einen Teil der Strukturinformationen mit sich.¹⁹⁹ Semistrukturierte Daten besitzen eine unregelmäßige bzw. partielle Struktur.²⁰⁰ Beispiele hierfür sind Extensible Markup Language-Daten (XML) oder Electronic Data Interchange-Dokumente (EDI).²⁰¹ Etwa achtzig Prozent der Daten eines Unternehmens sind unstrukturierte Daten.²⁰² Aus diesem Grund zählen insbesondere unstrukturierte Daten zum Gegenstandsbereich von ILM.²⁰³ Forschungsschwerpunkte in Bezug auf unstrukturierte Daten bilden Modelle zur Inhaltserschließung von Dokumenten,²⁰⁴ Methoden zur Ähnlichkeitsbestimmung von Anfragen zu Dokumenten und zur Indexierung von Dokumentenmengen,²⁰⁵ sowie die automatisierte Klassifikation von Dokumenten.²⁰⁶

¹⁹⁶ Vgl. Feldman, Sanger / Text Mining Handbook / 3-5

¹⁹⁷ Vgl. Bodrow, Bergmann / Wissensbewertung / 20-22; Gupta, Sharma, Hsu / Knowledge Based Organizations / 39-40

¹⁹⁸ Vgl. Piro, Gebauer / Datenarten / 146

¹⁹⁹ Vgl. Arimura / Patterns from Semi-structured Data / 2-3

²⁰⁰ Vgl. Abiteboul / Semi-Structured Data / 1-18

²⁰¹ Vgl. Loshin / Enterprise Knowledge Management / 65-67; Picot, Reichwald, Wigand / Unternehmung / 186-187

²⁰² Vgl. Gantz et al. / Digital Universe / 12; Nohr / Wissensmanagement / 7-16; Schoop / Informationsmanagement / 556-568

²⁰³ Vgl. Kanakamedala, Kaplan, Srinivasaraghavan / Data storage / 1-4

²⁰⁴ Vgl. Fugmann / Inhaltserschließung / 101-106

²⁰⁵ Vgl. Lepsky, Siepmann, Zimmermann / Automatische Indexierung / 47-56; Moens / Automatic Indexing / 343-347

²⁰⁶ Vgl. Chen / Information valuation / 135-146; Ellard et al. / File prediction / 1-14; Kaiser, Smolnik, Riempp / ILM-Framework / 483-484; Lee et al. / ILM with RFID / 112-113

2.4 Anforderungen an ILM

2.4.1 Klassifizierung der Informationen im Verlauf ihres Lebenszyklus

In Kapitel 2.1 wurde der Begriff Lebenszyklus bereits definiert. Im Zeitverlauf werden Informationen unterschiedlich oft verwendet bzw. weisen einen variierenden Wert für ein Unternehmen auf.²⁰⁷ In der Literatur²⁰⁸ werden verschiedene Kriterien für die Informationsklassifizierung verwendet. Denkbar ist die Klassifizierung der Informationen anhand der monetären Erträge, die aus der Nutzung der Informationen entstehen.²⁰⁹ Eine solche Klassifizierung erweist sich jedoch als sehr aufwändig, da sich aufgrund der Informationskomplexität ein direkter Bezug zu betriebswirtschaftlich relevanten Zielgrößen nur schwer herstellen lässt. Zudem gibt es in Unternehmen Informationen, welche zwar häufig genutzt werden, aber keine direkten monetären Erträge bewirken bzw. sich monetär kaum oder nur schwer bewerten lassen.²¹⁰ Es ist auch eine Klassifizierung von Informationen anhand der entstandenen Kosten möglich, die bei der Sammlung, Akquisition und Erstellung der Informationen anfallen.²¹¹ Je höher die entstandenen Kosten sind, desto höher ist der Wert, der den Informationen zugewiesen wird. Allerdings besteht die Möglichkeit, dass die Erzeugung der Informationen nur geringe Kosten verursacht hat, die Informationen aber dennoch einen hohen Wert für das Unternehmen haben.²¹² Die oben genannten Kriterien sind nicht für eine automatisierte Klassifizierung von Informationen geeignet, da sie sich nur schwer und mit hohem Aufwand bestimmen lassen und ihre Aussagekraft für das ILM gering ist. In dieser Arbeit wird der Nutzungsgrad als Kriterium für die Klassifizierung von Informationen herangezogen, welcher angibt, wie oft und wann der Zugriff auf eine Information erfolgte. Im

²⁰⁷ Vgl. Bhagwan et al. / Management of Data Storage /; Chen / Information valuation / 135-146; Shah et al. / Classification for ILM / 1-15

²⁰⁸ Die in Kapitel 2.4 erörterten Anforderungen an das ILM basieren auf der Literaturanalyse zum ILM. Vgl. hierzu Kapitel 1.3.

²⁰⁹ Vgl. Glazer / Measuring Information / 99-110; Moody, Walsh / Measuring Information / 496-512

²¹⁰ Vgl. Chen / Information valuation / 135-146; Shah et al. / Classification for ILM / 1-15; Al-Fedaghi / ILM / 336-337

²¹¹ Vgl. Moody, Walsh / Measuring Information / 496-512

²¹² Vgl. Oppenheim, Stenson, Wilson / Information as an asset / 458-464

Gegensatz zu den erstgenannten Kriterien lässt er sich ohne großen Aufwand ermitteln und automatisiert messen.²¹³

2.4.2 Technische und funktionale Anforderungen

Mit Hilfe von ILM werden Informationen, die in regelmäßigen Abständen genutzt werden, kostengünstig und direkt zugreifbar verwaltet.²¹⁴ Informationen, die einen hohen Nutzungsgrad²¹⁵ aufweisen, verbleiben im operativen System. Für die langfristige Aufbewahrung²¹⁶ eignen sich Informationen, die für längere Zeit nicht benötigt werden, aber aufgrund rechtlicher Anforderungen²¹⁷ aufbewahrt werden müssen. Aus technischer Sicht eignet sich eine dreistufige Speicherhierarchie²¹⁸ für das ILM. Diese ist in Abbildung 2-2 dargestellt.

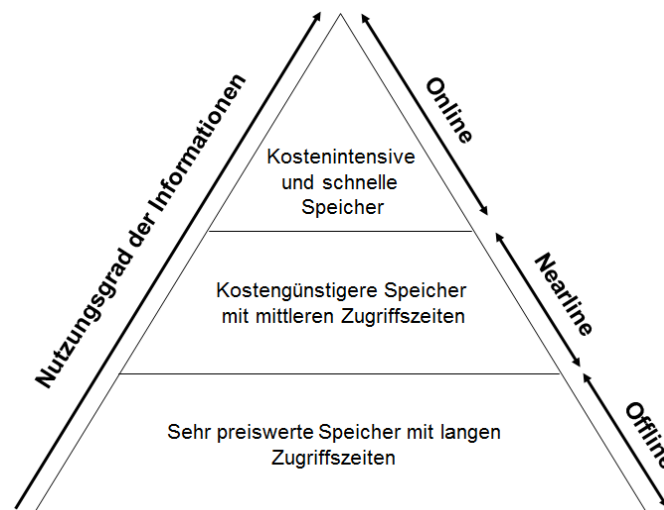


Abb. 2-2: Speicherhierarchie von ILM²¹⁹

²¹³ Vgl. Bhagwan et al. / Management of Data Storage /; Chen / Information valuation / 135-146; Chandra, Gehani, Yu / Automated Storage / 12; Douglass et al. / Delete-Optimized Storage /; Gavish, Sheng / Dynamic file migration / 177-189; Lenhardt et al. / Data Lifecycle / 1-4; Shah et al. / Classification for ILM / 1-15; Skyrme / Add value to business / 20-25

²¹⁴ Vgl. Kapitel 2.2

²¹⁵ Unter einem hohen Nutzungsgrad ist beispielsweise ein häufiger täglicher Zugriff auf die Daten zu verstehen.

²¹⁶ Vgl. Kapitel 2.1, Abbildung 2-1

²¹⁷ Vgl. Kapitel 2.4.3

²¹⁸ Vgl. Maier, Hädrich, Peinl / Knowledge Infrastructures / 254-257

²¹⁹ Vgl. nach Maier, Hädrich, Peinl / Knowledge Infrastructures / 255

Die Abbildung verdeutlicht den Sachverhalt, dass Informationen, die einen hohen Nutzungsgrad²²⁰ aufweisen, auf kostenintensiven Speichern mit kurzen Zugriffszeiten²²¹ abgelegt werden. Diese Speicher werden vorwiegend in den operativen Systemen eingesetzt. Die Informationen werden „Online“ gehalten. Auf der untersten Ebene befinden sich die für einen längeren Zeitraum nicht mehr benötigten Informationen, die „Offline“-Daten. Diese werden oft aufgrund rechtlicher Anforderungen langfristig aufbewahrt.²²² Hierfür eignen sich insbesondere die Daten, die archiviert werden.²²³ Die „Offline“-Daten können beispielsweise auf kostengünstige Magnetbänder²²⁴ verlagert werden. Die Speicherung auf Magnetband gilt als kostengünstigste Alternative bei sehr großen Datenbeständen.²²⁵ Denkbar ist in diesem Beispiel auch die Verwendung von WORM²²⁶-fähigen Magnetspeichern.²²⁷ Mit Hilfe der WORM-Technologie kann eine nachträgliche Änderung der Informationen im Sinne der rechtlichen Anforderungen verhindert werden. Der Zugriff auf diese Daten ist aufgrund der langen Zugriffszeiten²²⁸ mit einem hohen Zeitaufwand verbunden. Ein direkter Zugriff auf Archivdaten ist nicht möglich. Diese müssen zunächst in das operative System zurückgeladen werden.²²⁹ Im mittleren Bereich, in der Abbildung 2-2 als „Nearline“ gekennzeichnet, werden die Informationen abgelegt, auf welche in regelmäßigen Abständen zugegriffen wird, aber eine Ablage im operativen System nicht lohnend ist.²³⁰ Diese Ablagemöglichkeit wird

²²⁰ Vgl. Kapitel 2.4.1

²²¹ Die Zugriffszeit setzt sich zusammen aus der Zeit, die zur Positionierung des Schreib- bzw. Lesekopfes (seek time) benötigt wird und der Latenzzeit (latency), die vergeht, bis der richtige Sektor auf dem Speichermedium gefunden wurde (Vgl. hierzu Petrocelli / Information Lifecycle Management / 23). Die Zugriffszeiten liegen derzeit zwischen 2 Millisekunden (Festplatte) und 120 Sekunden (Magnetspeicher).

²²² Vgl. Kapitel 2.4.3

²²³ Werden Informationen „vollständig oder zumindest überwiegend in elektronischer, dem Rechner zugänglicher Form“ langfristig abgespeichert, so spricht man von Datenarchivierung (Vgl. hierzu Gulbins, Seyfried, Zimmermann / Archivierungssysteme / 10). „Langfristig“ kann dabei für einige Jahre, aber auch Jahrzehnte stehen (Vgl. hierzu Herbst / Archivierung in Datenbanksystemen / 194-211).

²²⁴ Vgl. Janko / Informationswirtschaft / 112-116

²²⁵ Vgl. Hillyer, Silberschatz / Storage Systems / 195-204

²²⁶ WORM steht für „Write once, read multiple times“ (Vgl. hierzu Janko / Informationswirtschaft / 126-127).

²²⁷ Vgl. Rüdiger / Stückwerk ILM / 19

²²⁸ Vgl. Härder, Rahm / Datenbanksysteme / 67-68

²²⁹ Vgl. Küspert, Schaarschmidt, Herbst / Archiv Transaktionen / 195-211

²³⁰ Vgl. Heinzel / Nearline Storage / 31

auch als Nearline-Storage (NLS) bezeichnet.²³¹ Möglich ist der Einsatz von Magnetplatten mit mittleren Zugriffszeiten.²³² Im Gegensatz zu den Informationen der operativen Systeme stellen zeitkritische Anfragen an Informationen der NLS-Ebene eher die Ausnahme dar.²³³ Vielmehr steht die Entlastung des kostenintensiven Speichers der operativen Systeme im Vordergrund. Die ILM-Software hat die Aufgabe, anhand vordefinierter Klassifizierungsfunktionen die Informationen entsprechend ihres Nutzungsgrads auf den Ebenen Online, Nearline oder Offline zu verwalten.²³⁴ Dabei muss die Anzahl der Datenverlagerungsvorgänge²³⁵ zwischen den verschiedenen Ebenen bzw. Speichermedien möglichst gering gehalten werden, da auch diese Kosten verursachen.²³⁶ Im Rahmen von ILM steht also nicht die Weiterentwicklung der Speichermedien im Vordergrund, sondern die Entwicklung eines ILM-Konzepts bzw. von ILM-Software, die die Informationen entsprechend ihres Nutzungsgrads automatisiert und kostengünstig in der Speicherhierarchie verwaltet.²³⁷ Bevor die Informationen auf die verschiedenen Speicherebenen verlagert werden, sollten diese bspw. in Dokumentenklassen eingeteilt werden.²³⁸ Durch eine Zuordnung der Daten zu Dokumentenklassen oder Dateiordnern werden diese als zusammengehörig betrachtet und klassifiziert. Somit kann vermieden werden, dass zusammengehörige Daten durch das ILM auf unterschiedlichen Speicherebenen abgelegt werden.²³⁹

²³¹ Vgl. Cao et al. / Grid Resource / 454 ; Liu, Li, Zhang / Dispatching Methods / 690-695; Maier, Hädrich, Peinl / Knowledge Infrastructures / 109

²³² Die Zugriffszeiten liegen in einem Bereich von 10-50 Sekunden (Vgl. hierzu Petrocelli / Information Lifecycle Management / 23; Rautenstrauch, Schulze / Informatik / 40).

²³³ Vgl. Gray, Reuter / Transaction Processing / 54

²³⁴ Wenn sich beispielsweise die Zugriffe auf zuvor auf NLS ausgelagerte Informationen deutlich erhöhen, muss die ILM-Software automatisch anhand der vordefinierten Regeln diese Informationen zurück zur Online-Ebene verlagern (Vgl. Maier, Hädrich, Peinl / Knowledge Infrastructures / 116; Mont / Privacy-aware ILM / 407; Doukas, Ntaikos, Bardis / ILM Applications / 492-497).

²³⁵ Vgl. Kapitel 5.1.2.4

²³⁶ Vgl. Lawrie, Randal; Barton / File Migration / 46-47

²³⁷ Der Nutzungsgrad der Informationen variiert im Zeitverlauf. Aus diesem Grund muss die Klassifizierung der Daten in definierten Zeitintervallen, bspw. monatlich, durchgeführt werden. Vgl. hierzu Turczyk / Wertzuweisung / 462.

²³⁸ Vgl. Kaiser, Smolnik, Riempp / ILM-Framework / 488

²³⁹ Problematisch wäre bspw., wenn Teile der Daten eines Geschäftsprozesses in der Online-Ebene verwaltet werden, und wiederum andere Teile archiviert oder gelöscht werden. Dies könnte unter Umständen die ordnungsgemäße Ausführung des Geschäftsprozesses verhindern. Vgl. hierzu Kaiser, Smolnik, Riempp / ILM-Framework / 487-489.

Abschließend soll das notwendige Zusammenspiel zwischen ILM-Software und Speicherhardware noch einmal aufgezeigt werden. Während Abbildung 2-1 einen stark vereinfachten Lebenszyklus von Informationen enthält, werden in Abbildung 2-3 beispielhaft auch mögliche Variationen des Nutzungsgrads²⁴⁰ während des Lebenszyklus für verschiedene Informationen dargestellt. Zusätzlich findet eine Einordnung der drei Ebenen „Online“, „Nearline“ und „Offline“ der Speicherhierarchie aus Abbildung 2-2 statt.

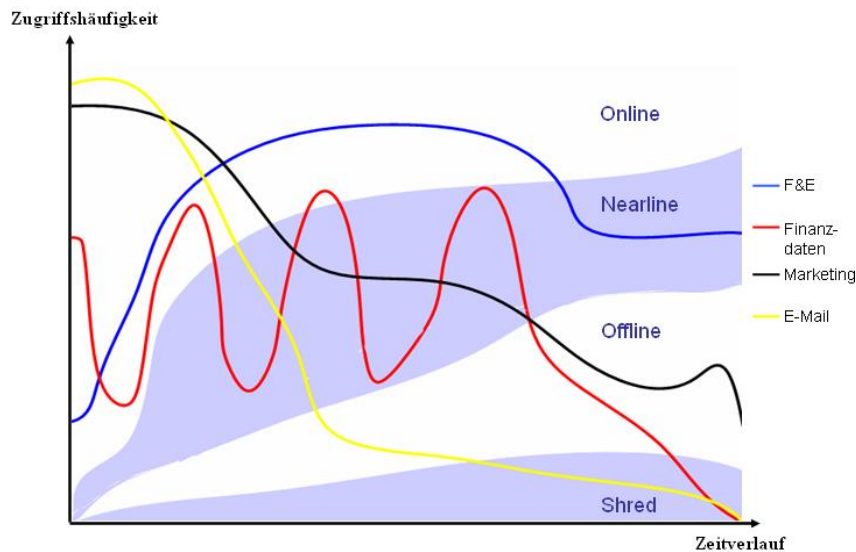


Abb. 2-3: Verwaltung der Daten im Verlauf ihres Lebenszyklus²⁴¹

Die Abbildung verdeutlicht, dass die verschiedenen Informationen zunächst im operativen System erstellt und hier für einen bestimmten Zeitraum „Online“ gehalten werden. Wie lange dabei bestimmte Informationen im operativen System verbleiben und ab welchem Nutzungsgrad die Informationen auf die Ebenen „Nearline“ und „Offline“ verlagert werden, wird in den zu definierenden Klassifizierungsfunktionen der ILM-Software vorgegeben. Wenn eine Aufbewahrung der Informationen nicht notwendig ist, wird ebenfalls definiert, wann die Löschung („Shred“) der Informationen erfolgen soll. Der Zugriff auf die Daten der Forschung und Entwicklung (F&E) nimmt im Laufe der Zeit zu. Als Beispiel sind verschiedene technische Baupläne für zukünftig zu entwickelnde Produkte zu nennen, die in Form von Computer-Aided-Design-Daten (CAD-Daten) vorliegen. Diese Baupläne können allerdings erst in Produkte umgesetzt werden, wenn

²⁴⁰ Diese Zugriffsänderungen während des Lebenszyklus werden durch die entsprechenden Amplitudenschwingungen dargestellt.

²⁴¹ In Anlehnung an Lundberg / Lifecycle Management / 5; Der Begriff „Shred“ steht für das Löschen der Daten.

die dafür benötigten Technologien vorhanden sind. Ist dies der Fall, erhöht sich der Zugriff auf die existierenden CAD-Daten. Wenn das Produkt nach einem bestimmten Zeitraum fertig gestellt wurde, sinkt automatisch der Zugriff auf die F&E-Daten. Damit werden diese auf die „Nearline“-Ebene im Sinne einer kostengünstigeren Datenhaltung auf die dafür vorgesehenen Speichermedien verlagert. Weiterhin ist in Abbildung 2-3 der typische Lebenszyklus von Finanzdaten dargestellt. Auf die Finanzdaten wird in der Regel periodisch zugegriffen, beispielsweise für die Erstellung des Quartalsabschlusses oder der Jahresbilanz.²⁴² Zeitweise erfolgen weniger Zugriffe, weshalb eine dauerhafte „Online“-Ablage nicht lohnend ist. Aus diesem Grund werden die Informationen nach kurzer Zeit auf die kostengünstigeren „Nearline“-Speicher verlagert.²⁴³ Sind die Finanzdaten für den Geschäftsablauf nicht mehr relevant, werden sie anschließend aufgrund der rechtlichen Anforderungen langfristig „Offline“ aufbewahrt. Hier empfiehlt sich die Archivierung der Informationen auf geeigneten Speichermedien, um die ordnungsgemäße Aufbewahrung zu gewährleisten.²⁴⁴ Nach Ablauf der gesetzlichen Aufbewahrungsfristen können die Finanzdaten gelöscht werden („Shred“).²⁴⁵ Marketingdaten wie Produktwerbungen oder Preisaktionen verlieren aufgrund der immer kürzer werdenden Produktlebenszyklen²⁴⁶ schnell an Bedeutung. Selbiges gilt für E-Mails. Nachdem diese gelesen wurden, erfolgt ein wiederholter Zugriff nur selten.²⁴⁷ Diese Informationen können anhand der vordefinierten Klassifizierungsfunktionen kurz nach ihrer Erstellung automatisch durch die ILM-Software auf kostengünstigere Speicher verlagert werden. Weiterhin sorgt eine rasche Archivierung oder Löschung der Informationen für die sichere Aufbewahrung bzw. Entlastung der kostenintensiven Speicher der operativen Systeme. Bevor die ILM-Software die verschiedenen Informationen automatisiert verwalten kann, ist die Implementierung von Klassifizierungsfunktionen erforderlich. Es muss zunächst von den ILM-Verantwortlichen in einem Unternehmen festgelegt werden, ab

²⁴² In Abbildung 2-3 erfolgt dabei eine stark vereinfachte Darstellung der Zugriffe auf die Finanzdaten in Form eines Lebenszyklus, der kontinuierlichen periodischen Schwankungen unterliegt.

²⁴³ Wenn sich der Zugriff auf die Informationen wieder erhöht, kann eine Verlagerung zurück in das operative System erfolgen. Dies muss anhand der vordefinierten Klassifizierungsfunktionen automatisch durch die ILM-Software gewährleistet werden.

²⁴⁴ Vgl. Kapitel 2.4.3

²⁴⁵ Vgl. Dauen / Aufbewahrungspflichten / 21-30

²⁴⁶ Vgl. Burmann, Freiling, Hülsmann / Ad-hoc-Krisen / 5; Kuhn, Hellingrath / SCM / 33

²⁴⁷ Selbstverständlich existieren in einem Unternehmen ebenfalls E-Mails, auf die für einen längeren Zeitraum häufig zugegriffen wird. Denkbar sind E-Mails, welche wichtige Kundenkontakte oder Vertragsdaten enthalten. Diese E-Mails verbleiben entsprechend länger auf der Online-Ebene.

welchem Nutzungsgrad bestimmte Informationen auf verschiedene Speichermedien durch die ILM-Software verlagert werden sollen. Dazu eignet sich die Erstellung eines unternehmensspezifischen ILM-Konzepts.²⁴⁸

2.4.3 Rechtliche Anforderungen

Nachdem funktionale und technische Anforderungen an ILM vorgestellt wurden, werden nun die in Deutschland existierenden rechtlichen Anforderungen an ILM aufgezeigt.²⁴⁹ Zu den rechtlichen Anforderungen zählen die Zulässigkeit verwendeter Speichermedien sowie die vollständige und richtige Übertragung bzw. der Transfer²⁵⁰ der Informationen. Der Schutz vor unbefugtem Zugriff sowie die geordnete Ablage müssen durch eine ordnungsgemäße Aufbewahrung der Informationen erfolgen. Für eine ordnungsgemäße Wiedergabe ist bezüglich der vollständigen Lesbarkeit der Informationen innerhalb der Aufbewahrungsfristen zu sorgen. Zusätzlich ist die Erstellung einer so genannten Verfahrensdokumentation erforderlich, um die Einhaltung der rechtlichen Vorschriften nachweisen zu können. Relevante Rechtsvorschriften, in Bezug auf das ILM, sind das Bundesdatenschutzgesetz (BDSG)²⁵¹, das Handelsgesetzbuch (HGB)²⁵², die Abgabenordnung (AO)²⁵³ sowie die aus diesen Rechtsvorschriften abgeleiteten Vorschriften der Grundsätze ordnungsgemäßer DV-gestützter Buchführungssysteme (GoBS)²⁵⁴ und die Grundsätze zum Datenzugriff und zur Prüfbarkeit digitaler Unterlagen (GDPdU)²⁵⁵. Für die Übertragung bzw. den Transfer, die Aufbewahrung und die Wiedergabe von Informationen werden nun die rechtlichen Anforderungen erläutert. Zusätzliche rechtliche Vorschriften und Anforderungen, die bei der Nutzung von Cloud-Computing-Diensten im ILM berücksichtigt werden müssen, werden in Kapitel 3.4.2 erörtert.

²⁴⁸ Vgl. Dannehl / Speicherbedarf / 8

²⁴⁹ Auf die Erörterung der rechtlichen Anforderungen in Bezug auf ausländische Rechtsvorschriften, wie beispielsweise der „Sarbanes Oxley Act“ in den USA, der „Data Protection Act“ in England oder das „Electronic Ledger Storage Law“ in Japan, wird im Rahmen dieser Arbeit verzichtet.

²⁵⁰ Beim Transfer findet die Übertragung der Informationen auf das Speichermedium statt.

²⁵¹ Vgl. Bundesministerium der Justiz / BDSG / §§ 1, 6, 35

²⁵² Vgl. Bundesministerium der Justiz / HGB / §§ 239, 257, 261

²⁵³ Vgl. Bundesministerium der Justiz / AO / §§ 146, 147, 239

²⁵⁴ Vgl. Bundesministerium der Finanzen / GoBS / 1-17

²⁵⁵ Vgl. Bundesministerium der Finanzen / GDPdU / 1-7

2.4.3.1 Rechtliche Anforderungen an die Übertragung bzw. den Transfer der Informationen

Für einen ordnungsgemäßen Transfer der Informationen bedarf es der Zulässigkeit des Speichermediums, auf das übertragen wird²⁵⁶ und einer vollständigen und richtigen Übertragung²⁵⁷. Für das ILM sind nach dem Gesetz alle Speichermedien zulässig, auf denen die Informationen gespeichert werden können.²⁵⁸ Um einen vollständigen und richtigen Transfer zu gewährleisten, müssen alle aufzubewahrenden Informationen ohne Übertragungsfehler auf das elektronische Speichermedium übertragen werden.

2.4.3.2 Rechtliche Anforderungen an die Ablage und Aufbewahrung der Informationen

Es existieren unterschiedliche Aufbewahrungsfristen für die in einem Unternehmen existierenden Informationen.²⁵⁹ Sämtliche Buchungsbelege müssen beispielsweise für zehn Jahre aufbewahrt werden.²⁶⁰ Die ordnungsgemäße Aufbewahrung der Informationen fordert den Schutz der Informationen vor unbefugtem Zugriff und die geordnete Ablage der Informationen.²⁶¹ Um den unbefugten Zugriff auf die Informationen zu verhindern, ist eine sichere Aufbewahrung unerlässlich. Sicherheit im Sinne der Informationsverarbeitung ist dabei ein Zustand, in dem alle sicherungswürdigen Informationen vor möglichen Beeinträchtigungen bewahrt sind.²⁶² Solche Beeinträchtigungen sind beispielsweise menschliches und technisches Versagen, Fehler bei der Informationsübertragung, Spionage oder höhere Gewalt.²⁶³ Um die Daten vor derartigen Beeinträchtigungen zu schützen, sollte ein Sicherheitskonzept erstellt und umgesetzt werden.²⁶⁴ Die so genannten Sicherheitsziele ergeben sich aus den sicherungswürdigen Sachverhal-

²⁵⁶ Vgl. Bundesministerium der Justiz / HGB / § 257 Abs. 3; Bundesministerium der Justiz / AO / § 147 Abs. 2; AWV / Aufbewahrungsverfahren / 9

²⁵⁷ Vgl. Bundesministerium der Justiz / HGB / § 239 Abs. 2; Bundesministerium der Justiz / AO / §§ 146 Abs. 1 und 239 Abs. 3

²⁵⁸ Vgl. Bundesministerium der Justiz / HGB / § 257 Abs. 3; Bundesministerium der Justiz / AO / § 147 Abs. 2. Es wird dabei vom Gesetzgeber keine Differenzierung vorgenommen, welche Informationen auf welchen Speichermedien abgelegt werden sollten.

²⁵⁹ Vgl. Dauen / Aufbewahrungspflichten / 33-122

²⁶⁰ Vgl. Dauen / Aufbewahrungspflichten / 52

²⁶¹ Vgl. AWV / Speicherplattensysteme / 10

²⁶² Vgl. Stelzer / Sicherheitsstrategien / 23

²⁶³ Vgl. BSI / Sicherheitshandbuch /

²⁶⁴ Vgl. Lippold, Stelzer, Konrad / Sicherheitskonzepte / 367-377

ten.²⁶⁵ Sicherungswürdige Informationen sind in dieser Arbeit unstrukturierte Daten, die im Rahmen des ILM verwaltet werden.²⁶⁶ Im Zusammenhang mit ILM stehen bezüglich der Sicherheitsziele die Integrität, die Vertraulichkeit sowie die langfristige Verfügbarkeit der Informationen im Vordergrund. Unter Integrität ist die Unverfälschtheit der Informationen zu verstehen.²⁶⁷ Eine nachträgliche Veränderung der Informationen darf nicht derart erfolgen, dass der ursprüngliche Inhalt nicht mehr feststellbar ist. Speichermedien, die nur einmal beschreibbar sind, bieten den Vorteil, dass gespeicherte Informationen bereits auf der Ebene des Speichermediums gegen nachträgliche Veränderungen geschützt werden können.²⁶⁸ Beim Einsatz wieder beschreibbarer Speichermedien darf eine Manipulation der Informationen oder der unbefugte Zugriff und die damit einhergehende Verletzung der Vertraulichkeit der Informationen nicht begünstigt werden.²⁶⁹ Handelt es sich bei den Informationen beispielsweise um personenbezogene Daten, so sind diese vor Missbrauch jeglicher Art zu schützen.²⁷⁰ Für das ILM ist weiterhin die Verfügbarkeit der ILM-Software sowie der eingesetzten Speichermedien im Sinne einer kurz-, mittel- und langfristigen Wiederverwendung²⁷¹ der Informationen von Bedeutung. Bezüglich der geordneten Ablage der Informationen bedarf es einer systematischen Ordnung der gespeicherten Informationen.²⁷² Die Art und Weise der Ordnung ist nicht näher festgelegt. Somit ist aus rechtlicher Sicht jede Ordnung akzeptabel, die ein systematisches Finden der Informationen in angemessener Zeit ermöglicht.²⁷³ Für die langfristige Aufbewahrung bzw. Archivierung der Informationen sind die eindeutige Kennzeichnung der Speichermedien und die Erstellung eines Archivverzeichnisses, in dem Daten, Erstellungsdatum und Aufbewahrungsfristen festgehalten werden, notwendig.²⁷⁴ Ein bestimmter Aufbewahrungsort ist gesetzlich nicht vorgeschrieben. Dieser

²⁶⁵ Vgl. Müller / IT-Sicherheit / 37-40

²⁶⁶ Vgl. Kapitel 2.3

²⁶⁷ Vgl. Kersten, Klett / IT-Security / 43-44

²⁶⁸ Vgl. Kapitel 2.4.2

²⁶⁹ Vgl. Bergmann / Speichermedien / 21-24

²⁷⁰ Vgl. Bundesministerium der Justiz / BDSG / § 1

²⁷¹ In welchen Zeitintervallen die Informationen verfügbar sein müssen, hängt insbesondere davon ab, ob die Informationen auf der Online, Nearline oder Offline-Ebene abgelegt wurden. Vgl. hierzu Kapitel 2.4.2.

²⁷² Vgl. Bundesministerium der Justiz / HGB / § 257 Abs.1 und § 257 Abs. 3

²⁷³ Vgl. AWV / Aufbewahrungsverfahren / 11

²⁷⁴ Vgl. Bundesministerium der Finanzen / GoBS / 13

muss lediglich im Geltungsbereich der Abgabenordnung, d.h. im Bereich der Bundesrepublik Deutschland liegen.²⁷⁵ Für Betriebsstätten außerhalb des Geltungsbereichs der Abgabenordnung ist entsprechend der dortigen Vorschriften zu verfahren.²⁷⁶

2.4.3.3 Rechtliche Anforderungen an die Wiedergabe der Informationen

Die lesbare Wiedergabe der Informationen wird ebenfalls per Gesetz gefordert.²⁷⁷ Sind zur Lesbarkeit besondere Hilfsmittel erforderlich, so sind diese auf Kosten des Unternehmens bereitzustellen.²⁷⁸ Bei der langfristigen elektronischen Aufbewahrung ist es notwendig, dass die eingesetzte ILM-Software sowie Speichermedien über die gesamte Dauer der Aufbewahrung zur Verfügung stehen. Im Falle eines Systemausfalls muss eine aktuelle Sicherheitskopie der gespeicherten Informationen aufbewahrt werden, um die Informationen bei Bedarf lesen zu können.²⁷⁹ Bei der Umstellung auf neue Speichermedien oder Speicherverfahren muss der alte Informationsbestand weiterhin lesbar bleiben. Ist dies nicht der Fall, ist der Unternehmer zur zusätzlichen Weiterführung des alten Systems bzw. Verfahrens solange verpflichtet, bis keine alten Informationsbestände mehr vorhanden sind.²⁸⁰ Insbesondere Informationen, die langfristig aufbewahrt werden, müssen im Rahmen einer angemessenen Frist lesbar gemacht werden.²⁸¹ Die „angemessene Frist“ ist gesetzlich nicht genau festgelegt. Es obliegt dem Betriebsprüfer, eine erfahrungsgemäße Frist in Anbetracht der jeweiligen Unternehmensverhältnisse zu setzen.²⁸²

2.4.3.4 Erstellung einer Verfahrensdokumentation

Um die Einhaltung der bereits erörterten rechtlichen Vorschriften bezüglich der Übertragung, der Ablage und Aufbewahrung sowie der Wiedergabe der Informationen nachzuweisen, ist eine Dokumentation der angewendeten Speicher- und Archivierungsver-

²⁷⁵ Vgl. Bundesministerium der Justiz / AO / § 146 Abs. 2

²⁷⁶ Eine Aufbewahrung von Buchführungsunterlagen außerhalb des Geltungsbereichs der Abgabenordnung ist nur nach vorheriger Genehmigung durch die zuständige Finanzbehörde möglich. Vgl. Bundesministerium der Justiz / AO / § 146 Abs. 2a.

²⁷⁷ Vgl. Bundesministerium der Justiz / HGB / §§ 239 Abs. 4, 257 Abs. 3 und 261

²⁷⁸ Vgl. Bundesministerium der Justiz / HGB / § 261

²⁷⁹ Vgl. AWV / Leitlinien / 26

²⁸⁰ Vgl. Bundesministerium der Finanzen / GoBS / 16, 17

²⁸¹ Vgl. Bundesministerium der Justiz / HGB / § 257 Abs. 3 Satz 2

²⁸² Vgl. AWV / Mikrofilm / 21

fahren zu erstellen. Zur Gestaltung der Verfahrensdokumentation existieren keine formalen Vorschriften.²⁸³ Inhaltlich muss sich diese aus einer Verfahrensbeschreibung und dem Nachweis über die ordnungsgemäße Durchführung der angewendeten Verfahren zusammensetzen.²⁸⁴ Im Rahmen der Verfahrensbeschreibung sollen organisatorische und technische Angaben zum Ablauf der Informationsablage und -aufbewahrung und der Informationswiederbeschaffung gemacht werden. Vor allem die programmtechnische Umsetzung sowie die Wahrung von Programm- und Datenintegrität sind zu beschreiben. Die Angaben zur Informationswiederbeschaffung müssen derart erfolgen, dass einem Dritten das Auffinden der Informationen kurzfristig möglich ist. Der Nachweis einer ordnungsgemäßen Aufbewahrung der Informationen kann durch die Erstellung entsprechender Protokolle unterstützt werden. Die Protokolle können automatisch oder manuell erstellt werden und sollten u.a. Angaben zu Informationsart, Transferzeitpunkt, Dateinamen und Indexierung machen.²⁸⁵ Die Verfahrensdokumentation ist für einen Zeitraum von zehn Jahren aufzubewahren.²⁸⁶

3 Cloud Computing

3.1 Begriff des Cloud Computings

Das Konzept des Cloud Computings stammt ursprünglich aus dem akademischen Umfeld. Technologischer Vorgänger des Cloud Computings ist das sogenannte Grid Computing, das seinen Ursprung im Gebiet der elektrischen Netze hat.²⁸⁷ Aufgrund der Nachfrage nach mehr Rechenleistung für wissenschaftliche Anwendungen wurden Hochleistungsrechner miteinander verbunden, um durch die gemeinsame Nutzung von Ressourcen große Datenmengen verwalten und evaluieren zu können. Für den kommerziellen Einsatz zur Verwaltung von Daten wurde der Begriff des Cloud Computings geprägt. Technologisch gesehen besteht der wesentliche Unterschied zwischen Grid Computing und Cloud Computing darin, dass im Cloud Computing der Zugriff auf die Ressourcen mit Hilfe von Internet-basierten Anwendungen erfolgt.²⁸⁸ Aus betriebswirt-

²⁸³ Vgl. Bundesministerium der Finanzen / GoBS / 14

²⁸⁴ Vgl. AWV / Verfahrensdokumentation / 11; Bundesministerium der Finanzen / GoBS / 14

²⁸⁵ Vgl. AWV / Verfahrensdokumentation / 17

²⁸⁶ Vgl. Bundesministerium der Justiz / HGB / § 257 Abs. 1-4 i. V. m. § 239 Abs. 2

²⁸⁷ Vgl. Dunkel et al./ Verteilte Anwendungen / 270; Vgl. Foster, Kesselmann, Tuecke / Anatomy of the Grid / 213; Vouk / Cloud Computing / 242-243

²⁸⁸ Vgl. Lawton / Moving OS to the Web / 16-19

schaftlicher Sicht wurden für das Cloud Computing zusätzlich Geschäftsmodelle entwickelt, um den kommerziellen Einsatz der Technologie in Unternehmen zu ermöglichen.²⁸⁹ Wesentlicher Treiber für den kommerziellen Einsatz waren Internetunternehmen, die große und nicht ausgelastete Rechenzentren besaßen. Der Online-Händler Amazon konnte bspw. seine IT-Infrastruktur lediglich im Weihnachtsgeschäft optimal auslasten, wenn die Anzahl der Anfragen und Einkäufe von Kunden anstieg. Die Verantwortlichen von Amazon entwickelten somit die Geschäftsidee, die ungenutzten Ressourcen ihrer Rechenzentren zu vermarkten. Die Ressourcen wurden daraufhin als Dienste über das Internet für andere Unternehmen und Nutzer angeboten.²⁹⁰

Der BITKOM versteht unter Cloud Computing „eine Form der bedarfsgerechten und flexiblen Nutzung von IT-Leistungen. Diese werden in Echtzeit als Service über das Internet bereitgestellt und nach Nutzung abgerechnet. Damit ermöglicht Cloud Computing eine Umverteilung von Investitions- zu Betriebsaufwand“²⁹¹. Das National Institute of Standards and Technology (NIST) definiert das Cloud Computing als verbrauchsorientierte, netzwerkbasierte, skalierbare und standortunabhängige Bereitstellung von Dienstleistungen.²⁹² Wenn eine Anwendung mehr Hardware-Ressourcen, wie Arbeitsspeicher oder Festplattenspeicher benötigt, können diese ohne großen Aufwand erweitert bzw. skaliert werden. Werden die Ressourcen nicht mehr benötigt, weil beispielsweise zeitlich schwankende Lastspitzen durch die Anwendung bedient wurden, können die Ressourcen ohne großen Aufwand wieder reduziert werden.²⁹³ Baun et al. definieren den Begriff Cloud Computing als bedarfsgerechte und flexible Bereitstellung von IT-Dienstleistungen über das Internet oder über ein Firmennetzwerk.²⁹⁴ Die Definition von Baun et al. wird dieser Arbeit zugrunde gelegt. Anbieter von Cloud-Computing-Diensten²⁹⁵ stellen IT-Dienstleistungen und IT-Infrastrukturen²⁹⁶ zur Verfügung, um

²⁸⁹ Vgl. Weiss / Computing in the Clouds / 16-25; Chang, Wills, De Roure / Cloud Business Models / 13-16

²⁹⁰ Vgl. Rimal, Choi, Lumb / Cloud Ecosystems / 21-22

²⁹¹ Vgl. BITKOM / Leitfaden Cloud Computing / 14

²⁹² Vgl. Mell, Grance / Definition Cloud Computing / 2

²⁹³ Vgl. Melzer / Architekturen mit Web Services / 174; Baun et al. / Cloud Computing / 1

²⁹⁴ Vgl. Dunkel et al. / Verteilte Anwendungen / 270; Baun et al. / Cloud Computing / 2

²⁹⁵ ILM hat sich in der Praxis bislang nicht durchgesetzt. Vgl. hierzu Thome, Sollbach / Modelle des ILM / 150; Baun et al. / Cloud Computing / 2. Im Gegensatz dazu stellen zahlreiche Anbieter von Cloud Computing mittlerweile Betriebssysteme, Anwendungssoftware oder Speicherplatz über das Internet als Dienst zur Verfügung. Vgl. hierzu Baun et al. / Cloud Computing / 23-38.

Ressourcen wie Hardware, Netzwerk und Speicherkapazitäten nutzungsabhängig bereitstellen und abzurechnen.²⁹⁷

3.2 Dimensionen des Cloud Computings

3.2.1 Organisatorische Dimension - Public, Private und Hybrid Cloud

Die organisatorische Dimension des Cloud Computings beschreibt das Zusammenspiel zwischen Anbietern und Nutzern von Cloud-Computing-Diensten. Es existieren drei verschiedene organisatorische Dimensionen, die Public-, Private- und Hybrid Cloud.²⁹⁸ Öffentlich zugängliche Dienste, die über das Internet genutzt werden können, werden von Anbietern einer Public Cloud angeboten. Die technische Infrastruktur befindet sich im Eigentum des Anbieters, die Nutzer greifen auf diese Infrastruktur zu und nutzen die bereitgestellten Dienste. Anbieter und Nutzer gehören unterschiedlichen Organisationen an. Die Abrechnung der Dienste erfolgt nutzungsabhängig.²⁹⁹ Auf den Ort der Datenhaltung und das verwendete Sicherheitskonzept³⁰⁰ hat der Nutzer in der Regel keinen Einfluss.³⁰¹ Die Public Cloud gilt als die kostengünstigste Alternative des Cloud Computings. Im Gegensatz zur Public Cloud befindet sich die technische Infrastruktur der zweiten organisatorischen Ebene, der Private Cloud, im Eigentum des Nutzers bzw. Unternehmens. Anbieter und Nutzer gehören derselben Organisation an. Der Zugriff auf die Dienste erfolgt über das Intranet. Die Gewährleistung der Datensicherheit und die bessere Kontrolle der Daten sind die Hauptgründe für die Entstehung von Private Clouds.³⁰² Durch die unternehmensinterne Bereitstellung von Infrastruktur und Diensten können diese besser auf die Geschäftsprozesse und Sicherheitsanforderungen eines Unternehmens angepasst werden, um eine standardisierte und sichere Umgebung zu gewährleisten.³⁰³ Durch die Nutzung und Administration der Private Cloud entstehen hö-

²⁹⁶ Unter dem Begriff IT-Infrastrukturen ist eine Ansammlung von Computern, Peripheriegeräten, Speichern und Netzwerkkomponenten zu verstehen. Vgl. hierzu Rumpel / IT-Infrastrukturen / 11.

²⁹⁷ Vgl. Baun et al. / Cloud Computing / 18

²⁹⁸ Vgl. Baun et al. / Cloud Computing / 15; Terplan, Voigt / Cloud Computing / 28-29

²⁹⁹ Anbieter von Public-Cloud-Diensten sind beispielsweise Microsoft, Google oder Amazon. Vgl. hierzu Armbrust et al. / Cloud Computing / 4.

³⁰⁰ Vgl. Kapitel 2.4.3.2

³⁰¹ Vgl. BITKOM / Leitfaden Cloud Computing / 30

³⁰² Vgl. Baun et al. / Cloud Computing / 15; Jamsa / Cloud Computing / 6

³⁰³ Vgl. Baun et al. / Cloud Computing / 15

here Kosten, als durch die Nutzung einer Public Cloud.³⁰⁴ Eine Mischform aus Public- und Private Cloud beschreibt die dritte organisatorische Dimension, die Hybrid Cloud.³⁰⁵ Durch die Kombination öffentlich zugänglicher und unternehmensinterner Dienste können die Vorteile beider organisatorischer Dimensionen genutzt werden. Denkbar ist, dass komplexe Analyseanwendungen durch Nutzung leistungsstarker und leicht zu skalierender Hardware-Ressourcen in einer Public Cloud ausgeführt werden. Anwendungen zur Verwaltung vertraulicher Daten, wie beispielsweise Kundendaten, Personaldaten oder Daten aus dem Bereich der Forschung und Entwicklung,³⁰⁶ können in der unternehmensinternen Private Cloud verbleiben.

3.2.2 Technische Dimension – Cloud-Computing-Dienste

Die technische Dimension beschreibt die angebotenen Dienste³⁰⁷, die von den Nutzern bzw. unterschiedlichen Zielgruppen des Cloud Computings genutzt werden. Wie in Abbildung 3-1 dargestellt, existieren drei technische Dimensionen: Infrastructure as a Service (IaaS), Platform as a Service (PaaS) und Software as a Service (SaaS).³⁰⁸

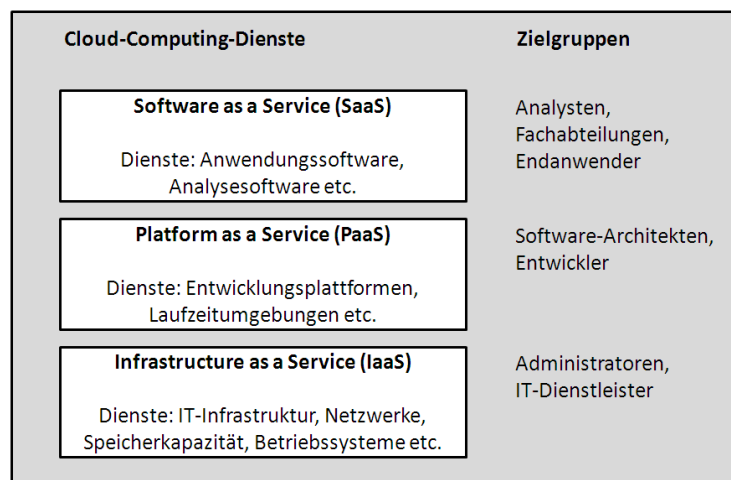


Abb. 3-1: Technische Dimensionen und Zielgruppen des Cloud Computings³⁰⁹

³⁰⁴ Vgl. Jamsa / Cloud Computing / 6

³⁰⁵ Vgl. Lin, Devine / Networks in Cloud Computing / 68

³⁰⁶ Vgl. Kapitel 2.4.2

³⁰⁷ Diese Dienste können über eine Private-, Public- oder Hybrid-Cloud bereitgestellt werden. Vgl. hierzu BITKOM / Leitfaden Cloud Computing / 28-31.

³⁰⁸ Vgl. Jin et al. / Cloud Types and Services / 346-349; BITKOM / Leitfaden Cloud Computing / 22-28

³⁰⁹ In Anlehnung an BITKOM / Leitfaden Cloud Computing / 23

Hardware, Netzwerke, Speicherkapazität, Betriebssysteme oder die gesamte IT-Infrastruktur werden durch Anbieter von IaaS bereitgestellt. Durch die Nutzung leistungsstarker Rechnerkapazitäten in einer Public Cloud können beispielsweise komplexe Datenanalysen durchgeführt werden.³¹⁰ Denkbar wäre auch die bedarfsgerechte Bereitstellung von Speicherplatzkapazität. Administratoren und IT-Dienstleister können dabei auf die technische Infrastruktur des IaaS-Anbieters zugreifen. Beispielsweise können Betreiber von Internetportalen bei einem starken Anstieg der Nutzerzahlen und Nutzeraktivitäten flexibel den benötigten Speicherplatz erweitern, wenn Nutzer neue Inhalte, wie Bilder, Videos oder Texte in das Internetportal einstellen. PaaS-Dienste richten sich an Software-Architekten und -entwickler.³¹¹ Über diese Dienste werden vom Anbieter Entwicklungsplattformen und Laufzeitumgebungen zur Verfügung gestellt, um Softwareanwendungen zu entwickeln und zu testen.³¹² Anbieter von SaaS-Diensten stellen Anwendungssoftware bereit. Endanwender können in einer Public Cloud über das Internet zum Beispiel auf CRM- oder ERP-Systeme zugreifen.³¹³

3.3 Unstrukturierte Daten im Kontext des Cloud Computings

Im Gegensatz zu strukturierten oder semistrukturierten Daten enthalten unstrukturierte Daten keine vorgegebenen Anordnungen von Datenfeldern oder Wertebereiche für Attribute.³¹⁴ Etwa achtzig Prozent der Daten eines Unternehmens sind unstrukturierte Daten.³¹⁵ Zu den unstrukturierten Daten zählen zum Beispiel E-Mails, Dokumente, Grafiken oder Videos.³¹⁶ Das Cloud Computing eignet sich für die Verwaltung unstrukturierter Daten.³¹⁷ In der Praxis nutzen Unternehmen das Cloud Computing bereits, um un-

³¹⁰ Das hätte den Vorteil, dass die Rechnerkapazitäten bedarfsgerecht bereitgestellt und genutzt werden können. Auf eine Anschaffung und der damit verbundenen Einrichtung und Wartung der Systeme durch Administratoren kann damit verzichtet werden (Vgl. hierzu Baun et al. / Cloud Computing / 3; McMillan / Role Cloud Computing in BigData /).

³¹¹ Vgl. McGrath / Understanding PaaS / 11-13

³¹² Vgl. Stankov, Datsenka / Global Software Development / 113-114

³¹³ Vgl. Grobman / ERP-Systeme On Demand / 15

³¹⁴ Vgl. Kapitel 2.3; Vgl. Feldman, Sanger / Text Mining Handbook / 3-5

³¹⁵ Vgl. Gantz et al. / Digital Universe / 12; Nohr / Wissensmanagement / 7-16; Schoop / Informationsmanagement / 556-568

³¹⁶ Vgl. Bodrow, Bergmann / Wissensbewertung / 20-22; Gupta, Sharma, Hsu / Knowledge Based Organizations / 39-40

³¹⁷ Vgl. Middleton / Data-Intensive Technologies for Cloud Computing / 90-92

strukturierte Daten zu verwalten.³¹⁸ Insbesondere die Nutzung von Speicherkapazitäten eines Public-Cloud-Dienstes eignet sich zur kostengünstigen Verwaltung unstrukturierter Daten.³¹⁹ Ein Unternehmen kann dabei über das Internet auf die Speichermedien des Cloud-Anbieters zugreifen. Bei der Nutzung von IaaS in einer Public Cloud entfallen die Kosten für die Administration unternehmensinterner Speichermedien. Die Abrechnung der Nutzungsgebühr erfolgt nutzungsabhängig und die Speicherkapazitäten können flexibel erhöht oder verringert werden.³²⁰

3.4 Anforderungen an das Cloud Computing

3.4.1 Technische und Funktionale Anforderungen

Mit Hilfe des Cloud Computings erfolgt die bedarfsgerechte und flexible Bereitstellung von IT-Dienstleistungen über das Internet oder über ein Firmennetzwerk.³²¹ Für den Nutzer von Cloud-Computing-Diensten bleibt die Komplexität der IT-Infrastruktur verborgen. Die Virtualisierung der Hard- und Softwareressourcen durch den Anbieter der Dienste ist zentraler Bestandteil des Cloud Computings.³²² Virtualisierung wird definiert als die Abstraktion logischer Systeme von der physischen Implementierung und Infrastruktur.³²³ Ein Beispiel ist die Virtualisierung von Speichermedien. Hierbei wird von den physischen Adressen eines Speichermediums abstrahiert, indem durch Softwarekomponenten virtuelle Adressen erzeugt werden.³²⁴ Die IT-Infrastruktur, bestehend aus Betriebssystemen, Speichermedien, Netzwerken und Anwendungsprogrammen, kann dem Nutzer mit Hilfe der Virtualisierung flexibel und bedarfsgerecht bereitgestellt werden.³²⁵ Neben der Virtualisierung sind die Konzepte der Webservices und der serviceorientierten Architekturen wesentliche Bestandteile des Cloud Computings. Webservices sind Dienste, die über Internet-basierte Protokolle in einer serviceorientierten Architektur genutzt werden können. Webservices dienen dem Aufruf, der Beschreibung

³¹⁸ Vgl. Baun et al. / Cloud Computing / 63-64

³¹⁹ Vgl. hierzu Talburt / Entity Resolution / 179-180; Gantz, Reinsel / Extracting Value / 2, 4, 11; Sriram / Cloud-Scale / 381-390; Mugg / Storage für KMU / 33. Administratoren und Anwender nutzen hierbei IaaS, bezogen auf die technische Ebene des Cloud Computings. Vgl. hierzu Kapitel 3.2.2.

³²⁰ Vgl. hierzu Kapitel 3.2.1 und Kapitel 3.2.2.

³²¹ Vgl. Dunkel et al. / Verteilte Anwendungen / 270; Baun et al. / Cloud Computing / 2

³²² Vgl. Baun, Kunze, Ludwig / Servervirtualisierung / 197-205

³²³ Vgl. BITKOM / Server-Virtualisierung / 4; Michelsen, English / Service Virtualization / 1-2

³²⁴ Vgl. Betzler / Virtualisierung / 102

³²⁵ Vgl. Kapitel 3.1

und dem Auffinden bzw. der Nutzung von Diensten.³²⁶ Das Konzept der serviceorientierten Architekturen unterstützt die Ausführung von Geschäftsprozessen durch fachliche Dienste. Die Services bzw. Dienste werden dabei von Anbietern bereitgestellt und haben einen klar definierten Umfang.³²⁷ Auf der untersten Ebene der Cloud-Computing-Architektur befinden sich die physischen Hard- und Softwareressourcen des Diensteanbieters. Durch die Virtualisierung können Cloud-Computing-Dienste, basierend auf den Hard- und Softwareressourcen, definiert und bereitgestellt werden. Die Nutzung dieser Dienste erfolgt über eine Nutzerschnittstelle. Der Nutzer greift mit Hilfe von Webservices und internetbasierten Anwendungen auf die Dienste zu. Dem Nutzer erscheint das Cloud-Computing-System als ein einziges System. Insbesondere bei Public Clouds sind die Systemumgebung und der angebotene Dienst jedoch meist auf verschiedene Rechner und Rechenzentren verteilt. Die Vorteile der verteilten Systeme sind die Bündelung und Steigerung der Rechenkapazität und -leistung sowie die Erhöhung der Ausfallsicherheit und Verfügbarkeit der Systeme.³²⁸ In Abbildung 3-2 ist der typische Aufbau einer Cloud-Computing-Architektur dargestellt.

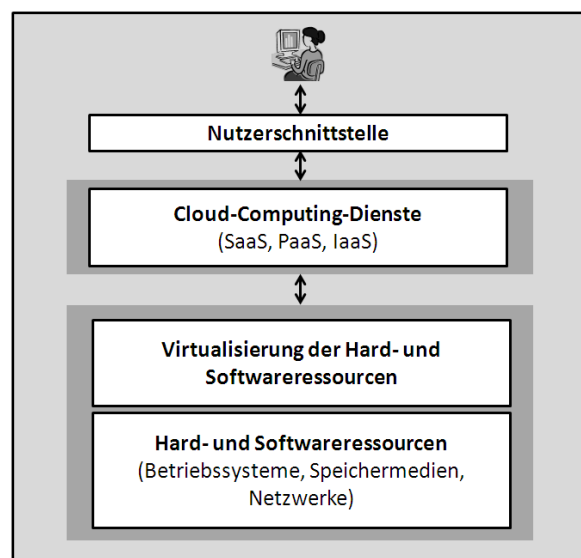


Abb. 3-2: Cloud-Computing-Architektur³²⁹

³²⁶ Vgl. Hansen, Neumann / Wirtschaftsinformatik / 802

³²⁷ Vgl. Beinhauer, Herr, Schmidt / SOA / 19-20; MacKenzie et al. / Reference Model SOA / 8-9

³²⁸ Vgl. Illik / Verteilte Systeme / 6-15

³²⁹ In Anlehnung an Cao, Li, Xia / Cloud Computing Architecture / 647-648

Aus technischer Sicht steht die Entwicklung von standardisierten Schnittstellen für den Datenaustausch zwischen verschiedenen Cloud-Computing-Diensten im Vordergrund.³³⁰ Es existieren zahlreiche Anbieter,³³¹ allerdings wurden bislang keine technischen Standards für das Cloud Computing entwickelt. Der Datenaustausch sowie ein Wechsel zwischen verschiedenen Cloud-Computing-Anbietern werden dadurch erschwert.³³² Der sogenannte „Anbieter-Lock-In“ beschreibt den Sachverhalt, dass ein Wechsel bzw. eine Migration der Daten zwischen unterschiedlichen Cloud-Computing-Anbietern nur schwer und zu hohen Kosten möglich ist.³³³ Aus diesem Grund sollte die Auswahl des Anbieters sehr sorgfältig erfolgen.

Die flächendeckende Durchsetzung von Breitbandinternetzugängen ermöglichte die Nutzung des Cloud Computings.³³⁴ Problematisch ist allerdings der Transfer großer Datenmengen in ein Cloud-System. Aufgrund geringer Upload-Geschwindigkeiten von Internetzugängen ist beim Transfer großer Datenmengen mit einem hohen Zeit- und Kostenaufwand zu rechnen.³³⁵ Diese Problematik können Nutzer des Cloud Computings umgehen. Hierzu können die physischen Speichermedien, auf denen die zu transferierenden Daten gespeichert sind, an den Cloud-Computing-Anbieter geschickt werden.³³⁶ Dieser transferiert anschließend die Daten in das Cloud-System. Der Anbieter berechnet auch den direkten Transfer der Daten von einem Speichermedium in das Cloud-System. Im Vergleich zu einem Transfer der Daten über das Internet stellt diese Alternative bei großen Datenmengen jedoch die kostengünstigere Variante dar.³³⁷ Zu beachten sind die Kosten für die Administration und den Versand der Datenträger, insbesondere wenn der Versand von Speichermedien auf wöchentlicher oder monatlicher Basis stattfindet. Anbieter von Cloud-Computing-Diensten stellen den Nutzern Werkzeuge zur Verfügung,

³³⁰ Vgl. Weinhardt et al. / Cloud-Computing / 460

³³¹ Vgl. Baun et al. / Cloud Computing / 2; Suter / Marktübersicht Cloud / 35-37; Schadler / Email in the Cloud / 14-15

³³² Vgl. Schnjakin, Meinel / Speicherressourcen in der Cloud / 2-3; Repschläger, Wind, Zarnekow / Anbieterauswahl in der Cloud / 2

³³³ Vgl. Höfer, Karagiannis / Cloud Computing Services / 87

³³⁴ Vgl. Mansmann / Breitband / 92

³³⁵ Vgl. Pamieri, Pardi / Cloud Computing / 131; Mahmood / Cloud Computing / 163-164

³³⁶ Vgl. Armbrust et al. / Cloud Computing / 16

³³⁷ Vgl. Yuan, Yang, Chen / Storage in the Cloud / 20

um die Dienste überwachen und verwalten zu können.³³⁸ Werkzeuge zur Überwachung der Dienste liefern zum Beispiel Informationen über den aktuellen Speicherverbrauch und die Auslastung der CPU.³³⁹ Weiterhin existieren Werkzeuge zur Netzwerk- und Speicherverwaltung sowie Zugriffs- und Berechtigungsverwaltung.³⁴⁰ Bei der Nutzung von Diensten über das Internet muss aus technischer Sicht die sichere Verarbeitung der Daten gewährleistet sein. Die Verschlüsselung der Daten ist hierbei von besonderer Bedeutung.³⁴¹ Nutzer sollten vor Vertragsabschluss überprüfen, inwieweit der Anbieter der Cloud-Computing-Dienste ein Konzept zur Verschlüsselung der Kommunikation und der verwalteten Daten implementiert hat.³⁴² Cloud-Anbieter können sich zertifizieren lassen, um sich die Einhaltung von Sicherheitsstandards bestätigen zu lassen.³⁴³

3.4.2 Rechtliche Anforderungen

Vom Gesetzgeber wird keine Differenzierung vorgenommen, welche Informationen auf welchen Speichermedien gespeichert werden sollen.³⁴⁴ Bei der Verwaltung unstrukturierter Daten mit Hilfe des Cloud Computings kommen zunächst die gesetzlichen Vorschriften zur Anwendung, die bereits in Kapitel 2.4.3 für das ILM erörtert wurden. Die vollständige und richtige Übertragung bzw. der Transfer³⁴⁵ der Informationen, der Schutz vor unbefugtem Zugriff sowie die ordnungsgemäße Aufbewahrung und Wiedergabe der Informationen zählen gleichermaßen zu den rechtlichen Anforderungen. Zusätzlich ist die Erstellung einer so genannten Verfahrensdokumentation erforderlich, um die Einhaltung der rechtlichen Vorschriften nachweisen zu können. Die für das Cloud

³³⁸ Vgl. Baun et al. / Cloud Computing / 41-47; Die vereinbarte Qualität und Menge der angebotenen Dienste, wie bspw. die Verfügbarkeit, Speicherkapazitäten, Zugriffsmöglichkeiten und Sicherung der Daten wird in den so genannten Service Level Agreements (SLA) festgehalten. Das SLA ist ein Vertrag, der zwischen Anbieter und Nutzer der Cloud-Computing-Dienste geschlossen wird (Vgl. hierzu Chang, Abu-Amara, Sanford / Cloud Services / 329-331).

³³⁹ Vgl. Cloud Climate / Watching the Clouds / ; CPU steht für "Central Processing Unit" (Vgl. Elliott / Information Technology / 370).

³⁴⁰ Vgl. Chang, Abu-Amara, Sanford / Cloud Services / 346; Baun et al. / Cloud Computing / 41-48

³⁴¹ Vgl. Mather, Kumaraswamy, Latif / Cloud Security / 61-63

³⁴² Vgl. BSI / Sicherheitsempfehlungen Cloud / 23-25

³⁴³ Nach erfolgreicher Überprüfung der Geschäftsprozesse des Cloud-Anbieters erhält dieser das ISO27001-Zertifikat. Vgl. Kersten, Reuter, Schröder / IT-Sicherheitsmanagement / 35-89; BSI / Sicherheitsempfehlungen Cloud / 25.

³⁴⁴ Vgl. Bundesministerium der Justiz / HGB / § 257 Abs. 3; Bundesministerium der Justiz / AO / § 147 Abs. 2.

³⁴⁵ Beim Transfer findet die Übertragung der Informationen auf die Cloud-Computing-Plattform statt.

Computing relevanten Rechtsvorschriften sind das Bundesdatenschutzgesetz (BDSG)³⁴⁶, das Handelsgesetzbuch (HGB)³⁴⁷, die Abgabenordnung (AO)³⁴⁸ und die aus diesen Rechtsvorschriften abgeleiteten Vorschriften der Grundsätze ordnungsgemäßer DV-gestützter Buchführungssysteme (GoBS)³⁴⁹ und die Grundsätze zum Datenzugriff und zur Prüfbarkeit digitaler Unterlagen (GDPdU)³⁵⁰. Bei der Nutzung von Cloud-Computing-Diensten über das Internet sind darüber hinaus weitere gesetzliche Vorschriften relevant. Diese werden im Folgenden ebenfalls erörtert.

3.4.2.1 Rechtliche Anforderungen an die Übertragung bzw. den Transfer der Informationen

Cloud-Dienste einer Public Cloud³⁵¹ werden über Telekommunikationsdienste erbracht. Beim Cloud Computing, insbesondere SaaS und PaaS³⁵² handelt sich um elektronische Informations- und Kommunikationsdienste bzw. Telemediendienste, da die Dienstleistung elektronisch über das Internet erbracht wird.³⁵³ Der Dienst Infrastructure-as-a-Service bzw. IaaS ist je nach Implementierungsform differenziert zu betrachten. Handelt es sich bei IaaS im Wesentlichen um die Verwaltung von IT-Infrastrukturen und Speichermedien, so werden keine Inhalte im Sinne von Informationsdiensten bereitgestellt. In diesem Fall ist der Anbieter von IaaS als Content Provider einzustufen.³⁵⁴ Da die Verwaltung der IT-Infrastruktur allerdings über das Internet erfolgt, kann IaaS auch als Telemediendienst eingestuft werden, wenn die Verwaltung der Hardwareressourcen und gespeicherten Informationen detailliert über Webservices und Web-Oberflächen möglich ist.³⁵⁵ Es existiert hierbei keine eindeutige Abgrenzung durch den Gesetzgeber.

³⁴⁶ Vgl. Bundesministerium der Justiz / BDSG / §§ 1, 6, 35

³⁴⁷ Vgl. Bundesministerium der Justiz / HGB / §§ 239, 257, 261

³⁴⁸ Vgl. Bundesministerium der Justiz / AO / §§ 146, 147, 239

³⁴⁹ Vgl. Bundesministerium der Finanzen / GoBS / 1-17

³⁵⁰ Vgl. Bundesministerium der Finanzen / GDPdU / 1-7

³⁵¹ Vgl. hierzu Kapitel 3.2.1. Beim Einsatz von Private Clouds sind entsprechend die rechtlichen Anforderungen und Vorschriften des ILM zu beachten, da in diesem Fall unternehmensinterne Ressourcen genutzt werden. Vgl. hierzu Kapitel 2.4.3 und Kapitel 3.2.

³⁵² Vgl. Kapitel 3.2.2

³⁵³ Vgl. Bundesministerium der Justiz / TMG / § 1, Absatz 1, Satz 1

³⁵⁴ Vgl. hierzu Bundesministerium der Justiz / TMG / § 10

³⁵⁵ Vgl. Bundesministerium der Justiz / TMG / § 1, Absatz 1, Satz 1

Der Transfer von Informationen über Telekommunikationsleitungen ist besonders schutzwürdig, da er Beeinträchtigungen wie Manipulation oder Spionage begünstigt.³⁵⁶ Die Übertragung bzw. der Transfer der Informationen ist deshalb durch geeignete Sicherheitsmaßnahmen, wie bspw. die elektronische Verschlüsselung, zu schützen.³⁵⁷ Somit findet Artikel 10 des Grundgesetzes (GG) Anwendung, wonach das Fernmeldegeheimnis nicht verletzt werden darf.³⁵⁸ Das Ausspionieren und Abhören von Informationen, die über Telemediendienste übertragen werden, ist verboten. Beschränkungen von Art. 10 GG dürfen nur durch den Erlass eines Gesetzes vorgenommen werden und wenn die Beschränkung dem Schutz der freiheitlichen demokratischen Grundordnung dient.³⁵⁹ Ergänzend hierzu definiert das Telekommunikationsgesetz (TKG), dass jeder Dienste-Anbieter zur Wahrung des Fernmeldegeheimnisses verpflichtet ist.³⁶⁰ Für einen ordnungsgemäßen Transfer der Informationen müssen die Informationen vollständig und richtig übertragen werden.³⁶¹ Für das Cloud Computing sind nach dem Gesetz alle Speichermedien zulässig, auf denen die Informationen gespeichert werden können.³⁶² Für einen vollständigen und richtigen Transfer müssen alle übertragenen Informationen ohne Übertragungsfehler auf die elektronischen Speichermedien der Cloud-Computing-Plattform übertragen werden.

3.4.2.2 Rechtliche Anforderungen an die Ablage und Aufbewahrung der Informationen

Die ordnungsgemäße Aufbewahrung der Informationen im Cloud Computing erfordert den Schutz der Informationen vor unbefugtem Zugriff und die geordnete Ablage der Informationen.³⁶³ Die für ILM relevanten Aufbewahrungsfristen gelten ebenso für das Cloud Computing.³⁶⁴ Die Informationen sind vor Beeinträchtigungen, wie menschliches

³⁵⁶ Vgl. Venhaus, Haselbeck, Wintermann / Schatten der Überwachung / 40-50; Wicker / Death of Privacy / 129-134

³⁵⁷ Vgl. Kapitel 3.4.1

³⁵⁸ Vgl. Bundesministerium der Justiz / GG / Art. 10, Abs. 1

³⁵⁹ Vgl. Bundesministerium der Justiz / GG / Art. 10, Abs. 2

³⁶⁰ Vgl. Bundesministerium der Justiz / TKG / §88

³⁶¹ Vgl. Bundesministerium der Justiz / HGB / § 239 Abs. 2; Bundesministerium der Justiz / AO / §§ 146 Abs. 1 und 239 Abs. 3

³⁶² Vgl. Bundesministerium der Justiz / HGB / § 257 Abs. 3; Bundesministerium der Justiz / AO / § 147 Abs. 2.

³⁶³ Vgl. AWV / Speicherplattensysteme / 10

³⁶⁴ Vgl. Dauen / Aufbewahrungspflichten / 33-122; Vgl. Kapitel 2.4.3.2

und technisches Versagen, Fehler bei der Informationsübertragung, Spionage oder höhere Gewalt, zu schützen.³⁶⁵ Hierzu müssen die Nutzer und Anbieter von Cloud-Computing-Diensten Sicherheitskonzepte³⁶⁶ erstellen und umsetzen. Aufgrund der verteilten Architektur von Cloud-Computing-Plattformen sind diese meist über verschiedene Rechenzentren und Server verteilt.³⁶⁷ Dies erschwert den Schutz der Informationen vor Beeinträchtigungen.³⁶⁸ Es genügt nicht, wenn nur der Anbieter der Dienste ein entsprechendes Sicherheitskonzept umsetzt. Durch das Zusammenspiel zwischen Nutzer und Anbieter, der genutzten Client-Rechner auf Nutzerseite sowie der verwendeten Kommunikationskanäle, muss auch der Nutzer der Dienste für den Schutz vor Beeinträchtigungen sorgen.³⁶⁹

In Bezug auf die Sicherheitsziele stehen im Cloud Computing die Vertraulichkeit, Transparenz und Integrität im Vordergrund. Die Manipulation von Informationen oder der unbefugte Zugriff auf die im Cloud Computing verwalteten Informationen dürfen im Sinne der Vertraulichkeit nicht begünstigt werden.³⁷⁰ Handelt es sich bei den Informationen beispielsweise um personenbezogene Daten, so sind diese vor Missbrauch jeglicher Art zu schützen.³⁷¹ Zur Sicherstellung der Leistung und angemessenen Nutzung ihrer IT-Infrastruktur halten Anbieter von Public-Cloud-Diensten die gespeicherten Informationen ihrer Kunden redundant oder transferieren diese oft in unterschiedliche Rechenzentren.³⁷² Dabei werden die Informationen oft auf Systemen von Subunternehmen des Anbieters gespeichert und es werden Landesgrenzen überschritten.³⁷³ Hierin liegt eine wesentliche Problematik des Cloud Computings hinsichtlich der Transparenz der Datenverarbeitung. Durch den Einsatz verschiedener Algorithmen und Dateisysteme in unterschiedlichen Rechenzentren und bei Subunternehmen kann der Cloud-Anbieter nicht immer eine durchgängige Verschlüsselung der Informationen garantie-

³⁶⁵ Vgl. BSI / Sicherheitshandbuch / ; Vgl. Kapitel 3.4.1

³⁶⁶ Vgl. Lippold, Stelzer, Konrad / Sicherheitskonzepte / 367-377

³⁶⁷ Vgl. Kapitel 3.4.1

³⁶⁸ Vgl. Winkler / Securing the Cloud / 55-66

³⁶⁹ Vgl. Tsvihun, Stephanow, Streitberger / Sicherheit Cloud Computing / 18-19

³⁷⁰ Vgl. Bergmann / Speichermedien / 21-24

³⁷¹ Vgl. Bundesministerium der Justiz / BDSG / § 1

³⁷² Vgl. Bauer, Adams / Service Quality Cloud / 93-95

³⁷³ Vgl. Buyya, Broberg, Goscinski / Cloud Principles / 603-605

ren.³⁷⁴ Bezüglich des Aufbewahrungsortes der Informationen geben die meisten Cloud-Anbieter keine Zusicherungen darüber, in welchem Land die Informationen gespeichert werden und welche Maßnahmen ergriffen werden, um die Vertraulichkeit zu gewährleisten.³⁷⁵ Ein genauer Aufbewahrungsort ist gesetzlich zwar nicht vorgeschrieben. Dieser muss lediglich im Geltungsbereich der Abgabenordnung, im Bereich der Bundesrepublik Deutschland liegen.³⁷⁶ Eine Aufbewahrung von Buchführungsunterlagen in einer Public Cloud außerhalb des Geltungsbereichs der Abgabenordnung ist nur nach vorheriger Genehmigung durch die zuständige Finanzbehörde zulässig.³⁷⁷ Für Betriebsstätten außerhalb des Geltungsbereichs der Abgabenordnung ist entsprechend der dortigen Vorschriften zu verfahren. Bevor sich die Verantwortlichen eines Unternehmens dazu entschließen, Informationen mittels Public-Cloud-Diensten zu verwalten, sollten die Aspekte der Vertraulichkeit, und welche Zusicherungen der Cloud-Anbieter in der SLA macht, genau überprüft werden. Bei Unwissenheit über die Einhaltung von Sicherheitsstandards oder über den Ablageort der Informationen verstoßen die Verantwortlichen gegen geltendes deutsches Recht. Ein weiteres Problem ist, wenn bspw. Informationen einer Public Cloud in den USA gespeichert werden bzw. der Anbieter seinen Firmensitz in den USA hat. Die National Security Agency (NSA) nutzt die Daten von Unternehmen wie Microsoft, Google, Yahoo oder Facebook, die Dienste zur Verwaltung von Informationen in Public-Cloud-Systemen anbieten, zu Analyse- und Auswertungszwecken.³⁷⁸ Eine Unterscheidung zwischen vertraulichen und nicht vertraulichen Daten wird dabei nicht vorgenommen. Die Integrität³⁷⁹ bzw. die Unverfälschtheit der Informationen ist ebenfalls zu gewährleisten.³⁸⁰ Die Informationen dürfen im Cloud Computing nicht unbemerkt geändert und gelöscht werden. Hierzu sind entsprechende Sicherheitsvorkehrungen durch Nutzer und Anbieter von Cloud-Diensten zu treffen, wie die Verschlüsselung der Daten und der Notwendigkeit der Authentifizierung der Nutzer.³⁸¹

³⁷⁴ Vgl. Kriesi / Rechtliche Aspekte von SaaS / 40-41; Linthicum / Cloud Contracts /

³⁷⁵ Vgl. Kriesi / Rechtliche Aspekte von SaaS / 40-41; Gantz, Reinsel / Extracting Value / 9

³⁷⁶ Vgl. Bundesministerium der Justiz / AO / § 146 Abs. 2

³⁷⁷ Vgl. Bundesministerium der Justiz / AO / § 146 Abs. 2a

³⁷⁸ Vgl. Levy / NSA killed the Internet / ; Venhaus, Haselbeck, Wintermann / Schatten der Überwachung / 40-50; Wicker / Death of Privacy / 129-134

³⁷⁹ Vgl. Kersten, Klett / IT-Security / 43-44

³⁸⁰ Vgl. Kersten, Klett / IT-Security / 43-44

³⁸¹ Vgl. Li et al. / Authentication Cloud / 157-166

3.4.2.3 Rechtliche Anforderungen an die Wiedergabe der Informationen

Werden Informationen im Cloud Computing verwaltet, so wird vom Gesetz die lesbare Wiedergabe der Informationen gefordert.³⁸² Wenn zur Lesbarkeit besondere Hilfsmittel erforderlich sind, müssen diese auf Kosten des nutzenden Unternehmens bereitgestellt werden.³⁸³ In der Regel erfolgt der Zugriff und die Wiedergabe der Informationen mit Hilfe von Webservices und internetbasierten Anwendungen, die vom Anbieter der Dienste bereitgestellt werden und für die der Nutzer eine entsprechende Gebühr entrichtet.³⁸⁴ Bei der langfristigen elektronischen Aufbewahrung ist es notwendig, dass die genutzten Cloud-Computing-Dienste über die gesamte Dauer der Aufbewahrung zur Verfügung stehen. Die Verfügbarkeit der Cloud-Dienste und die redundante Datenhaltung werden in den SLA zwischen Nutzer und Anbieter vereinbart.³⁸⁵ Sollten aus bestimmten Gründen³⁸⁶ Informationen nach dem Transfer von unternehmenseigenen Systemen in eine Public Cloud nicht mehr lesbar sein, ist der Unternehmer verpflichtet, das alte System weiterzuführen.³⁸⁷ Insbesondere Informationen, die langfristig aufbewahrt werden, müssen innerhalb einer angemessenen Frist lesbar gemacht werden.³⁸⁸ Der Betriebsprüfer legt dabei eine erfahrungsgemäße Frist, in Anbetracht der jeweiligen Unternehmensverhältnisse, fest.³⁸⁹

3.4.2.4 Erstellung einer Verfahrensdokumentation

Zum Nachweis der Einhaltung der erörterten rechtlichen Vorschriften im Cloud Computing ist eine Verfahrensdokumentation zu erstellen. Es existieren keine gesetzlichen Vorgaben zur formalen Gestaltung der Verfahrensdokumentation.³⁹⁰ Inhaltlich muss sich diese aus einer Verfahrensbeschreibung und dem Nachweis über die ordnungsgemäße Durchführung der angewendeten Verfahren zusammensetzen.³⁹¹ Bei der Verfah-

³⁸² Vgl. Bundesministerium der Justiz / HGB / §§ 239 Abs. 4, 257 Abs. 3 und 261

³⁸³ Vgl. Bundesministerium der Justiz / HGB / § 261

³⁸⁴ Vgl. Kapitel 3.4.1

³⁸⁵ Vgl. Chang, Abu-Amara, Sanford / Cloud Services / 329-331

³⁸⁶ Denkbar sind Daten, die nur im Zusammenhang mit bestimmter Anwendungssoftware lesbar sind.

³⁸⁷ Vgl. Bundesministerium der Finanzen / GoBS / 16, 17

³⁸⁸ Vgl. Bundesministerium der Justiz / HGB / § 257 Abs. 3 Satz 2

³⁸⁹ Vgl. AWV / Mikrofilm / 21

³⁹⁰ Vgl. Bundesministerium der Finanzen / GoBS / 14

³⁹¹ Vgl. AWV / Verfahrensdokumentation / 11; Bundesministerium der Finanzen / GoBS / 14

rensbeschreibung sollen organisatorische und technische Angaben zum Ablauf bei der Ablage, Aufbewahrung und Wiedergabe der Informationen gemacht werden. Der Anbieter der Cloud-Computing-Dienste sollte sämtliche Vorgänge dokumentieren und auf Anfrage des Nutzers Logdateien und Protokolle bereitstellen können. Da insbesondere bei der Nutzung einer Public Cloud die IT-Infrastruktur durch den Anbieter verwaltet wird,³⁹² ist der Nutzer weitestgehend auf den Anbieter angewiesen.³⁹³ Von daher sollte auch die ordnungsgemäße Dokumentation der Vorgänge durch den Anbieter in der SLA vereinbart werden. Die Protokolle können automatisch oder manuell erstellt werden und sollten u.a. Angaben zu Informationsart, Transferzeitpunkt, Dateinamen und Indexierung machen.³⁹⁴ Die Verfahrensdokumentation ist für einen Zeitraum von zehn Jahren aufzubewahren.³⁹⁵

4 Wirtschaftlichkeitsanalyse

4.1 Begriff der Wirtschaftlichkeitsanalyse

Die Wirtschaftlichkeit beschreibt das Verhältnis von Kosten und Nutzen.³⁹⁶ Die Wirtschaftlichkeitsanalyse umfasst die Analyse der Kosten- und Nutzenstruktur sowie die Analyse der Beziehungen zwischen Kosten und Nutzen.³⁹⁷ Der Begriff Kosten³⁹⁸ wird definiert als der in Geldeinheiten bewertete sachzielbezogene Güterverzehr bzw. Ressourcenverbrauch innerhalb einer Periode.³⁹⁹ Direkte Kosten stellen bspw. Kosten für Hardware, Software und Personal dar. Direkte Kosten können einem Kostenträger di-

³⁹² Vgl. Kapitel 3.2.1

³⁹³ Vgl. Wang et al. / Cloud Methodology / 414

³⁹⁴ Vgl. AWV / Verfahrensdokumentation / 17

³⁹⁵ Vgl. Bundesministerium der Justiz / HGB / § 257 Abs. 1-4 i. V. m. § 239 Abs. 2

³⁹⁶ Vgl. Rau / Statistik / 246-248; Peters, Brühl, Stelling / Betriebswirtschaftslehre / 209

³⁹⁷ Vgl. hierzu Wieczorrek, Mertens / Management IT-Projekte / 206; Heinrich, Stelzer / Informationsmanagement / 363-366. Mit Hilfe der Wirtschaftlichkeitsanalyse können Investitionen im IT-Bereich, Systeme oder Systementwürfe, Produkte und Dienstleistungen sowie die Entwicklung bzw. der Einsatz von Methoden und Werkzeugen, beurteilt werden. Vgl. hierzu Heinrich, Stelzer / Informationsmanagement / 363-366.

³⁹⁸ Die Kosten werden in Kostenarten gegliedert. Es werden vier Kostenarten unterschieden: Verbrauchsgebundene Kosten wie Energiekosten, betriebsgebundene Kosten wie Personal- und Wartungskosten, kapitalgebundene Kosten wie Zinsen und Abschreibungen sowie sonstige Kosten wie Steuern und Versicherungskosten. Vgl. hierzu VDI / Betriebswirtschaftliche Berechnungen / 5-6 ; Schröder / Operatives Controlling / 24.

³⁹⁹ Vgl. Schmidt / Wirtschaftlichkeit / 262; Heinen / Führungslehre / 313

rekt zugeordnet werden.⁴⁰⁰ Indirekte Kosten sind Gemeinkosten. Beispiele für indirekte Kosten sind Raumkosten oder Verwaltungskosten. Der Begriff Nutzen wird definiert als die Fähigkeit eines Gutes, zur Befriedigung eines Bedürfnisses beizutragen.⁴⁰¹ Verschiedene Nutzenkategorien sind der direkt quantifizierbare Nutzen sowie der indirekt bzw. schwer- oder nicht quantifizierbare Nutzen.⁴⁰² Durch die Analyse der Beziehungen zwischen Kosten und Nutzen können die Beziehungszusammenhänge zwischen den einzelnen Kosten- und Nutzenarten untersucht werden.⁴⁰³ Zur Ergänzung der Wirtschaftlichkeitsanalyse können die Methoden der Wirtschaftlichkeitsrechnung angewendet werden.⁴⁰⁴ Mit Hilfe der Methoden der Wirtschaftlichkeitsrechnung werden Indikatoren der Wirtschaftlichkeit, wie bspw. die Rentabilität oder die Amortisationsdauer, berechnet.⁴⁰⁵

4.2 Methoden der Wirtschaftlichkeitsrechnung

4.2.1 Ermittlungsmodelle

Die sogenannten Ermittlungsmodelle der Wirtschaftlichkeitsrechnung basieren auf monetären Größen bzw. Kennzahlen.⁴⁰⁶ Voraussetzung für die Anwendung von Ermittlungsmodellen ist die quantitative Messbarkeit der zu betrachtenden Kennzahlen.⁴⁰⁷ Bezüglich der Ermittlungsmodelle wird zwischen statischen und dynamischen Methoden unterschieden. Statische Methoden berücksichtigen im Gegensatz zu den dynamischen Methoden keine zeitlichen Aspekte bei der Entstehung von Einzahlungen oder

⁴⁰⁰ Vgl. Preißler / Kosten- und Leistungsrechnung / 36

⁴⁰¹ Vgl. Antweiler / Wirtschaftlichkeitsanalyse / 98ff.; Bechmann / Nutzwertanalyse / 48

⁴⁰² Vgl. Möhrle, Isenmann / Zukunftsstrategien / 158; Heinen / Führungslehre / 314; Direkter Nutzen entsteht bspw. durch die Senkung von Material- oder Personalkosten. Durch Verringerung des Lagerbestands können die Lagerkosten gesenkt werden, wodurch ein indirekter Nutzen entsteht. Nicht monetär messbarer Nutzen entsteht bspw. durch die Verbesserung des Informationsangebots in einem Unternehmen und den damit verbundenen Auswirkungen auf die Qualität von Entscheidungen. Vgl. hierzu Heinrich, Stelzer / Informationsmanagement / 366; Böhm, Fuchs / System-Entwicklung / 85-86.

⁴⁰³ Es wird untersucht, welche Kosten welchen Nutzen erzeugen oder umgekehrt, welche Nutzenart welche Kosten verursacht. Dabei kann zusätzlich überprüft werden, inwiefern bspw. die Veränderung der Kostenstruktur zu einer Erhöhung des Nutzens führt. Vgl. hierzu Fritz / Ökonomischer Nutzen / 146-147; Zimmermann, Henke, Broer / Finanzwissenschaft / 98-100.

⁴⁰⁴ Vgl. Becker, Schütte / Handelsinformationssysteme / 188-189; Voegelé, Sommer / Wirtschaftlichkeitsrechnung / 343; Heinrich, Stelzer / Informationsmanagement / 362

⁴⁰⁵ Vgl. Kapitel 4.2

⁴⁰⁶ Vgl. Krieger / Kostenrechnung / 59

⁴⁰⁷ Vgl. Meyer / Betriebswirtschaftliche Kennzahlen / 17-25

Auszahlungen monetärer Größen.⁴⁰⁸ Im Folgenden werden die statischen und dynamischen Methoden der Wirtschaftlichkeitsrechnung vorgestellt.

4.2.1.1 Statische Methoden

Zu den statischen Methoden der Wirtschaftlichkeitsrechnung gehören die Kosten- und Gewinnvergleichsrechnung, die Rentabilitätsrechnung und die Amortisationsrechnung.⁴⁰⁹ Mit Hilfe der Kostenvergleichsrechnung kann aus mehreren zur Verfügung stehenden Alternativen die kostengünstigste⁴¹⁰ Investition ermittelt werden.⁴¹¹ Voraussetzung für die Anwendung der Kostenvergleichsrechnung ist, dass die zu vergleichenden Investitionen den gleichen Zweck erfüllen und gleich hohe Erlöse bzw. den gleichen Nutzen erzielen.⁴¹² Typische Vergleichsmerkmale von Investitionen sind die Nutzungsdauer, die Leistungsmenge sowie das investierte Kapital.⁴¹³ Die Gewinnvergleichsrechnung stellt eine Erweiterung der Kostenvergleichsrechnung dar.⁴¹⁴ Ihre Anwendung eignet sich, wenn die betrachteten Investitionen mit unterschiedlichen Erlösen verbunden sind.⁴¹⁵ Aufbauend auf den Ergebnissen der Kosten- und Gewinnvergleichsrechnung kann eine Rentabilitätsrechnung durchgeführt werden. Ziel der Rentabilitätsrechnung ist es, das Verhältnis von Gewinn und eingesetztem Kapital, bezogen auf die betrachteten Investitionen, zu ermitteln.⁴¹⁶ Bei der Rentabilität von Investitionen wird zwischen der Kapitalrentabilität und der Umsatzrentabilität unterschieden.⁴¹⁷ Die Kapitalrentabilität drückt das Verhältnis von Gewinn und eingesetztem Kapital aus, die Umsatzrentabilität stellt den auf den Umsatz bezogenen Gewinnanteil dar.⁴¹⁸ Um den Zeit-

⁴⁰⁸ Vgl. Horst / Investition / 37

⁴⁰⁹ Vgl. Tauberger / Controlling / 95; Rau / Statistik / 205

⁴¹⁰ Die betrachteten Kosten können Durchschnittskosten der betrachteten Investition bezogen auf eine bestimmte Laufzeit oder die Kosten eines bestimmten Jahres sein. Vgl. hierzu Rau / Statistik / 207.

⁴¹¹ Vgl. Becker / Finanzwirtschaft / 45

⁴¹² Vgl. Rau / Statistik / 207

⁴¹³ Vgl. Röhrich / Investitionsrechnung / 11-20; Grundsätzlich sind alle während der Nutzungsdauer entstandenen Kosten wie Entwicklungs-, Betriebs- und Instandhaltungskosten in die Berechnung einzubeziehen. Vgl. hierzu Voegelé, Sommer / Wirtschaftlichkeitsrechnung / 346.

⁴¹⁴ Vgl. Becker / Finanzwirtschaft / 51

⁴¹⁵ Vgl. Poggensee / Investitionsrechnung / 66-67

⁴¹⁶ Vgl. Voegelé, Sommer / Wirtschaftlichkeitsrechnung / 353

⁴¹⁷ Vgl. Obermeier, Gasper / Investitionsrechnung / 19

⁴¹⁸ Vgl. Bieg, Kußmaul / Finanzierung / 4; Hering / Taschenbuch Wirtschaftsingenieure / 363-364; Matschke, Hering, Klingelhöfer / Finanzplanung / 3-4

raum zu ermitteln, in dem das investierte Kapital wieder zurück in das Unternehmen fließt, wird eine Amortisationsrechnung durchgeführt.⁴¹⁹ Die Investition mit der kürzesten Amortisationsdauer ist vorzuziehen.⁴²⁰ Die statischen Methoden der Wirtschaftlichkeitsrechnung stellen Kosten, Gewinn, Rentabilität und Amortisationsdauer der zu vergleichenden Investitionen gegenüber. Durch Anwendung der statischen Methoden kann sowohl das Auswahlproblem⁴²¹ von Investitionen als auch das Ersatzproblem⁴²² von Investitionen beurteilt werden.⁴²³ Da die statischen Methoden zeitliche Aspekte bei der Entstehung von Einzahlungen oder Auszahlungen von monetären Größen nicht berücksichtigen, sollten zusätzliche Überlegungen, wie Preis-, Lohn- oder Zinsentwicklung, in die Investitionsrechnungen einfließen.⁴²⁴

4.2.1.2 Dynamische Methoden

Im Gegensatz zu den statischen Methoden berücksichtigen die dynamischen Methoden die Zeitpunkte bei der Entstehung von Einzahlungen oder Auszahlungen monetärer Größen.⁴²⁵ Die dynamischen Methoden der Wirtschaftlichkeitsrechnung basieren auf Ein- und Auszahlungsströmen.⁴²⁶ Es werden die Zeitpunkte der für eine Investition relevanten Ein- und Auszahlungen in die Betrachtungen einbezogen.⁴²⁷ Durch Anwendung der dynamischen Methoden kann die Vorteilhaftigkeit, das Auswahlproblem so-

⁴¹⁹ Vgl. Jung / Controlling / 21; Hölscher / Investition / 33-34

⁴²⁰ Vgl. Obermeier, Gasper / Investitionsrechnung / 35; Die Amortisationsrechnung dient lediglich als ergänzende Entscheidungshilfe. In die Berechnung der Amortisationsdauer sollten weitere Überlegungen einfließen, um die Vorteilhaftigkeit einer Investition zu ermitteln. Führt eine Investition bspw. zu einer Senkung von Personalkosten und damit zu einer kurzen Amortisationsdauer, so sollte zusätzlich überprüft werden, welche Auswirkungen durch die Einsparung von Personal auf andere betriebliche Bereiche, außerhalb der betrachteten Investition, entstehen. Vgl. Hoffmeister / Investitionsrechnung / 140.

⁴²¹ Bei Vorliegen mehrerer Investitionsalternativen wird die vorteilhafteste Investition ermittelt. Vgl. hierzu Zimmermann, Fries, Hoch / Rechnungswesen / 276.

⁴²² Existiert bereits bspw. eine betriebliche Anlage, so wird untersucht, ob und wann die bestehende Anlage durch eine neue Anlage aus Wirtschaftlichkeitsgesichtspunkten ersetzt werden soll. Vgl. Becker / Finanzwirtschaft / 39.

⁴²³ Vgl. Wöltje / Investition und Finanzierung / 83

⁴²⁴ Vgl. Zimmermann, Fries, Hoch / Rechnungswesen / 284-291; Hierzu können bspw. ergänzend Trend- und Sensitivitätsanalysen durchgeführt werden. Vgl. hierzu Pfnür, Schetter, Schöbener / Risikomanagement / 76-79; Hering / Investitionstheorie / 308-319.

⁴²⁵ Vgl. Schmeisser et al. / Finanzierung und Investition / 55

⁴²⁶ Vgl. Carstensen / Investitionsrechnung / 31

⁴²⁷ Ein- und Auszahlungen können nach Höhe und Zeitpunkt unterschiedlich sein. Deshalb werden die Zahlungen auf den Zeitpunkt unmittelbar vor Investitionsbeginn abgezinst. Die Abzinsung ist das zentrale Merkmal der dynamischen Methoden der Wirtschaftlichkeitsrechnung. Vgl. hierzu Becker / Wirtschaftswissen / 178.

wie die optimale Laufzeit einer Investition untersucht werden.⁴²⁸ Die Vorteilhaftigkeit drückt aus, ob eine Investition lohnend ist. Ob eine Investition einer alternativen Investition vorzuziehen ist, wird mit dem sogenannten Auswahlproblem untersucht. Die optimale Laufzeit einer Investition liefert Aussagen dazu, nach welchem Zeitraum eine getätigte Investition durch eine neue Investition ersetzt werden sollte. Zu den dynamischen Methoden der Wirtschaftlichkeitsrechnung gehören die Kapitalwertmethode, die Annuitätenmethode, die dynamische Amortisationsrechnung und die interne Zinsfußmethode.⁴²⁹ Bei der Kapitalwertmethode wird zur Beurteilung einer Investition der Kapitalwert herangezogen.⁴³⁰ Der Kapitalwert ergibt sich aus der Summe aller auf einen bestimmten Zeitpunkt⁴³¹ ab- oder auf gezinsten⁴³² positiven und negativen Zahlungen, die der betrachteten Investition zugeordnet werden können.⁴³³ Eine Investition wird als vorteilhaft angesehen, wenn der Kapitalwert größer als der Wert Null ist.⁴³⁴ Existieren zwei Investitionen, so ist die Investition mit dem größeren Kapitalwert vorzuziehen.⁴³⁵ Die Annuitätenmethode stellt eine Abwandlung der Kapitalwertmethode dar.⁴³⁶ Nach Ermittlung des Kapitalwerts wird der durchschnittliche jährliche Überschuss, die Annuität, berechnet.⁴³⁷ Eine Investition ist dann lohnend, wenn der durchschnittliche jährliche Überschuss nicht negativ ist.⁴³⁸ Ziel der dynamischen Amortisationsrechnung ist die Ermittlung der Zeitspanne, in der die Auszahlungen für eine Investition durch späte-

⁴²⁸ Vgl. Wöltje / Investition und Finanzierung / 131-144; Vgl. Schmeisser et al. / Finanzierung und Investition / 59-74

⁴²⁹ Vgl. Schierenbeck, Wöhle / Betriebswirtschaftslehre / 410

⁴³⁰ Obermeier, Gasper / Investitionsrechnung / 71-74

⁴³¹ Hierfür wird als Bezugszeitpunkt der Zeitpunkt vor der Investition gewählt. Vgl. hierzu Becker / Wirtschaftswissen / 178.

⁴³² Der verwendete Zinssatz wird als Kalkulationszinsfuß bezeichnet. Vgl. hierzu Rolfes / Investitionsrechnung / 22-23.

⁴³³ Die Kapitalwertmethode gibt also an, welchen monetären Überschuss eine Investition erwirtschaftet. Vgl. hierzu Becker / Wirtschaftswissen / 178.

⁴³⁴ Vgl. Ventzislavova, Hensel / Betriebswirtschaftliche Formelsammlung / 40

⁴³⁵ Vgl. Hering / Investitionstheorie / 118; Becker / Wirtschaftswissen / 179

⁴³⁶ Vgl. Peters, Brühl, Stelling / Betriebswirtschaftslehre / 103

⁴³⁷ Vgl. Poggensee / Investitionsrechnung / 145; Die Annuität betrachtet den periodisch konstanten Überschuss einer Investition.

⁴³⁸ Vgl. Poggensee / Investitionsrechnung / 145

re Einzahlungen zurückgewonnen werden.⁴³⁹ Eine Investition ist dann vorteilhaft, wenn ihre dynamische Amortisationsdauer eine vorgegebene maximale Amortisationsdauer nicht überschreitet.⁴⁴⁰ Die jährliche Rendite einer Investition wird mit der internen Zinsfußmethode berechnet.⁴⁴¹ Diese Methode stellt eine Umkehrung der Kapitalwertmethode dar.⁴⁴² Es wird der Zinssatz berechnet, der zu einem Kapitalwert von Null führt.⁴⁴³ Der ermittelte interne Zinsfuß gibt die erwartete Effektivverzinsung für die betrachtete Investition an.⁴⁴⁴ Die dynamischen Methoden der Wirtschaftlichkeitsrechnung berücksichtigen Ein- und Auszahlungen sowie die relevanten Zahlungszeitpunkte.⁴⁴⁵ Zins- und Zinseszinsseffekte werden durch die Zuordnung von Ein- und Auszahlungen zu den jeweiligen Zeitpunkten erfasst.⁴⁴⁶ Die Annahmen zu künftigen Ein- und Auszahlungen basieren auf Schätzungen bzw. Prognosen, weshalb die Anwendung dynamischer Methoden mit einer gewissen Unsicherheit verbunden ist.⁴⁴⁷ Weiterhin ist der Planungsaufwand im Vergleich zu den statischen Methoden höher, da die Zahlungsströme für die einzelnen Zahlungszeitpunkte bestimmt werden müssen.⁴⁴⁸

⁴³⁹ Vgl. Kußmaul / Betriebswirtschaftslehre / 224; Zusätzlich wird untersucht, ob durch die Einzahlungsüberschüsse auch die Verzinsung des gebundenen Kapitals zum Kalkulationszinssatz zurückgewonnen werden kann. Im Gegensatz zur statischen Amortisationsrechnung (Vgl. hierzu Kapitel 4.2.1.1) werden bei der dynamischen Amortisationsrechnung Zins- und Zinseszinsseffekte berücksichtigt.

⁴⁴⁰ Die maximal vorgegebene Amortisationsdauer darf dabei die Nutzungsdauer der Investition nicht übersteigen. Vgl. hierzu Kußmaul / Betriebswirtschaftslehre / 224.

⁴⁴¹ Vgl. Becker / Wirtschaftswissen / 180

⁴⁴² Vgl. Gugel / Investitionsplanung / 15; Baier / Controlling / 249

⁴⁴³ Vgl. Jung / Controlling / 130

⁴⁴⁴ Vgl. Peters, Brühl, Stelling / Betriebswirtschaftslehre / 104

⁴⁴⁵ Im Gegensatz dazu berücksichtigen die statischen Methoden Kosten und Erträge. Zeitliche Aspekte werden nicht berücksichtigt. Vgl. hierzu Kapitel 4.2.1.1.

⁴⁴⁶ Der Kapitalwert einer Investition mit anfangs höheren und später geringeren Kapitalrückflüssen ist höher als der Kapitalwert einer Investition mit anfangs geringeren und später höheren Kapitalrückflüssen. Vgl. hierzu Obermeier, Gasper / Investitionsrechnung / 77.

⁴⁴⁷ Vgl. Andree / Wirtschaftlichkeitsanalyse / 150; Röhrich / Investitionsrechnung / 109-112

⁴⁴⁸ Damit sich der Aufwand lohnt, eignen sich die dynamischen Methoden hauptsächlich zur Beurteilung von größeren Investitionen. Vgl. hierzu Becker / Wirtschaftswissen / 180.

4.2.2 Quantitative Entscheidungsmodelle

Entscheidungsmodelle dienen der Untersuchung der Vorteilhaftigkeit mehrerer zur Auswahl stehender Entscheidungen. Eine Entscheidung⁴⁴⁹ stellt die Auswahl einer oder mehrerer Alternativen durch einen Entscheidungsträger dar.⁴⁵⁰ Die Entscheidung ist dabei an ein bestimmtes Ziel des Entscheidungsträgers gekoppelt.⁴⁵¹ Quantitative Entscheidungsmodelle beschreiben die vorhandenen Alternativen durch die Zuordnung von Zahlenwerten.⁴⁵² Ziel quantitativer Methoden ist die Analyse betrieblicher und wirtschaftlicher Prozesse zur Entscheidungsvorbereitung.⁴⁵³ Dazu werden mathematische Methoden angewendet.⁴⁵⁴ Es wird zwischen analytischen Methoden und Simulationen unterschieden.⁴⁵⁵

4.2.2.1 Analytische Methoden

Durch Anwendung analytischer Methoden wird ein zu lösendes Realproblem in ein mathematisches Problem, das sogenannte Formalproblem, übertragen.⁴⁵⁶ Neben der modellhaften⁴⁵⁷ Strukturierung des Realitätsausschnitts erfolgt die Formulierung eines mathematischen Modells.⁴⁵⁸ Zum mathematischen Modell wird anschließend ein Zielope-

⁴⁴⁹ Die Entscheidungstheorie ist eine wissenschaftliche Disziplin zur Behandlung von Entscheidungen. Diese untergliedert sich in die normative und deskriptive Entscheidungstheorie. Die normative Entscheidungstheorie beschäftigt sich mit der Frage, wie ein Entscheidungsträger, unter Berücksichtigung eines definierten Ziels, die vorteilhafteste Alternative ermitteln kann (Vgl. hierzu Homburg / Quantitative Betriebswirtschaftslehre / 455-456). Hierbei werden Richtlinien vorgegeben, wie sich ein Entscheidungsträger in bestimmten Entscheidungssituationen verhalten soll. Vgl. hierzu Hagenloch / Entscheidungslehre / 2. Die Beschreibung tatsächlicher Entscheidungen in konkreten Situationen erfolgt in der deskriptiven Entscheidungstheorie. Im Vordergrund steht das tatsächliche Problemlösungsverhalten eines Entscheidungsträgers. Vgl. hierzu Homburg / Quantitative Betriebswirtschaftslehre / 455; Stelling / Controlling / 313.

⁴⁵⁰ Vgl. Homburg / Quantitative Betriebswirtschaftslehre / 455

⁴⁵¹ Vgl. Hagenloch / Entscheidungslehre / 2; Fink / Strategische Unternehmensberatung / 79-80

⁴⁵² Quantitative Daten lassen sich einfacher erheben, als qualitative Daten. Vgl. hierzu Kapitel 4.2.3; Riesenhuber / Empirische Forschung / 7.

⁴⁵³ Vgl. Homburg / Quantitative Betriebswirtschaftslehre / 42

⁴⁵⁴ Vgl. Gohout / Operations Research / 3

⁴⁵⁵ Vgl. Heinrich, Stelzer / Informationsmanagement / 369; Foscht, Angerer, Swoboda / Mixed Methods / 252

⁴⁵⁶ Vgl. Biethahn, Mucksch, Ruf / Informationsmanagement / 162-163; Heinrich, Stelzer / Informationsmanagement / 369

⁴⁵⁷ Ein Modell wird definiert als die formalisierte Beschreibung der realen Situation. Durch Anwendung bestimmter Funktionen werden die relevanten Variablen und ihre Beziehungen zueinander repräsentiert. Vgl. hierzu Grochla / Organisationstheorie / 85.

⁴⁵⁸ Vgl. Stötzer / Performance Reporting / 344

rator⁴⁵⁹ hinzugefügt, wodurch das Formalproblem entsteht.⁴⁶⁰ Entsprechend der Zielsetzung erfolgt die analytische Ermittlung einer Lösung.⁴⁶¹ Zur Problemlösung kommen mathematische Methoden zum Einsatz. Beispiele für mathematische Methoden sind die statische Programmierung⁴⁶², die dynamische Programmierung⁴⁶³, Entscheidungsbaumverfahren⁴⁶⁴, Netzplantechnik⁴⁶⁵, Warteschlangenmodelle⁴⁶⁶ und spieltheoretische Modelle⁴⁶⁷.⁴⁶⁸ Anschließend wird die gefundene Lösung auf die Realität übertragen.⁴⁶⁹ Das entwickelte Modell und die ermittelte Lösung behalten solange ihre Gültigkeit, bis sich wesentliche Parameter der Realität ändern.⁴⁷⁰ Tritt dieser Fall ein, müssen die Modellparameter des Modells angepasst und eine neue Lösung hergeleitet werden.⁴⁷¹

4.2.2.2 Simulationenmethoden

Der Einsatz von Simulationenmethoden eignet sich zur Nachbildung komplexer Ereignisse oder Systeme, um Berechnungen oder Untersuchungen durchzuführen.⁴⁷² Die Simulation ist ein heuristisches Verfahren, bei dem durch Experimentieren am Modell kom-

⁴⁵⁹ Es wird festgelegt, für wen und für welchen Zweck modelliert wird. Vgl. hierzu Biethahn, Mucksch, Ruf / Informationsmanagement / 163.

⁴⁶⁰ Vgl. Zelewski, Hohmann, Hügens / Planungs- und Steuerungssysteme / 344-345

⁴⁶¹ Vgl. House / Evaluating / 80-81

⁴⁶² Bei der statischen Programmierung bleiben die Variablen und Funktionen über den betrachteten Zeitraum hinweg konstant. Es wird unterschieden zwischen der linearen, nicht linearen und ganzzahligen Programmierung. Vgl. hierzu Zimmermann, Stache / Quantitative Methoden / 48-125, 208-211; Thonemann / Operations Management / 546.

⁴⁶³ Die dynamische Programmierung wird angewendet, wenn sich das betrachtete Problem in mehrere Zustände aufteilen lässt. Es wird die Abhängigkeit eines Zustands mit nachfolgenden Zuständen im Zeitverlauf untersucht. Vgl. hierzu Stepan, Fischer / Optimierung / 282-285.

⁴⁶⁴ Die zu erarbeitenden Lösungen werden in einem Entscheidungsbaum dargestellt. So können entweder alle möglichen Lösungen erarbeitet oder bestimmte Gruppen von Lösungen ausgeschlossen werden. Vgl. hierzu Rau / Statistik / 86-87.

⁴⁶⁵ Netzpläne bilden die Abhängigkeiten zwischen verschiedenen Ereignissen ab. Vgl. hierzu Biethahn, Mucksch, Ruf / Informationsmanagement / 400-404.

⁴⁶⁶ Warteschlangenmodelle beschäftigen sich mit der Analyse der Wartezeiten von Ereignissen, bspw. wenn ein Ereignis auf die Abarbeitung oder Bedienung warten muss. Vgl. hierzu Vahrenkamp / Produktionsmanagement / 91-93.

⁴⁶⁷ Die Zusammenhänge und Abhängigkeiten zwischen Entscheidungen und Akteuren werden mit spieltheoretischen Modellen gelöst. Hierbei wird untersucht, ob bestimmte Entscheidungen miteinander konkurrieren. Vgl. hierzu Miebach / Handlungstheorie / 404-406.

⁴⁶⁸ Vgl. Kummer, Grün, Jammerneegg / Produktion und Logistik / 287

⁴⁶⁹ Vgl. Daume / Finanzmathematik / 178

⁴⁷⁰ Vgl. Biethahn, Mucksch, Ruf / Informationsmanagement / 164

⁴⁷¹ Vgl. Peters / Organisationsgestaltung / 19-20

⁴⁷² Vgl. Kolonko / Stochastische Simulation / 1-2

plexe wirtschaftliche oder technische Sachverhalte untersucht werden.⁴⁷³ Es wird zwischen deterministischen und stochastischen Simulationsmethoden unterschieden.⁴⁷⁴ Die deterministische Simulation eignet sich für die Analyse von Problemen, bei denen alle eingehenden Daten bekannt sind.⁴⁷⁵ Durch Experimente können Aussagen über organisatorische Bearbeitungsfolgen getroffen werden.⁴⁷⁶ Die Untersuchung von Problemen, bei denen zufällige Ereignisse⁴⁷⁷ im Vordergrund stehen, ist Gegenstand der stochastischen Simulation.⁴⁷⁸ Zur Modellierung und Nachbildung zufälliger Ereignisse werden Zufallszahlen generiert.⁴⁷⁹ Um statistisch signifikante Aussagen über das betrachtete Problem oder System liefern zu können, müssen mehrere Simulationsläufe durchgeführt werden.⁴⁸⁰ Neben den geringen Kosten für die Untersuchung komplexer Probleme stellt die Flexibilität einen weiteren Vorteil simulationsbasierter Methoden dar.⁴⁸¹ Das Simulationsmodell kann beliebig verändert werden, um damit verbundene Auswirkungen auf die Simulationsergebnisse zu untersuchen. Somit können Rückschlüsse auf das Verhalten des realen Systems gezogen werden.⁴⁸² Hierfür eignen sich insbesondere die Methoden der stochastischen Simulation.⁴⁸³

⁴⁷³ Vgl. Zimmermann, Stache / Quantitative Methoden / 336; Banks / Simulation / 3-4; Velten / Modeling and Simulation / 7

⁴⁷⁴ Vgl. Zimmermann, Stache / Quantitative Methoden / 337; Peters / Organisationsgestaltung / 109

⁴⁷⁵ Beispielsweise zur Simulation der Lagerhaltung oder bei der Erstellung von Fahr- und Schichtplänen. Vgl. hierzu Zimmermann, Stache / Quantitative Methoden / 337; Eichholz, Vilkner / Wirtschaftsmathematik / 371.

⁴⁷⁶ Es werden verschiedene Alternativen untersucht, an denen Veränderungen vorgenommen werden können. Vgl. hierzu Nebl / Produktionswirtschaft / 513 und Kapitel 5.2.1.

⁴⁷⁷ Der Einfluss der stochastischen bzw. zufälligen Ereignisse auf das betrachtete Gesamtsystem kann nicht vollständig analytisch erfasst werden. Beispiele für zufällige Ereignisse sind Verkehrsabläufe im Straßenverkehr oder Niederschlagsmengen an einem bestimmten Ort. Vgl. hierzu Schmidt, Klüver, Klüver / Soft Computing / 63; Henze / Stochastik / 105.

⁴⁷⁸ Vgl. Kolonko / Stochastische Simulation / 1-2

⁴⁷⁹ Vgl. hierzu Hübner / Stochastik / 163-165; Kapitel 5.2.2.1.3 zur Vorgehensweise bei der Erzeugung von Zufallszahlen.

⁴⁸⁰ Vgl. hierzu Eichholz, Vilkner / Wirtschaftsmathematik / 373; Fujimoto, Perumalla, Riley / Network Simulation / 37. Zur Sicherstellung statistischer Signifikanz eignet sich die Durchführung von dreißig Simulationsläufen. Vgl. hierzu Sachs, Hedderich / Angewandte Statistik / 308-309; Nissen, Biethahn / Stochastische Optimierung / 117; McPeak / Simulation Noise / 364-366; Railsback, Grimm / Agent-Based Modeling / 197.

⁴⁸¹ Vgl. Han, Qiu, Huang / Reference Framework / 86

⁴⁸² Vgl. Zimmermann, Stache / Quantitative Methoden / 336

⁴⁸³ Vgl. Heinrich, Stelzer / Informationsmanagement / 369; Kapitel 5.2.1

4.2.3 Qualitative Entscheidungsmodelle

Die Ermittlungsmodelle und quantitativen Entscheidungsmodelle der Wirtschaftlichkeitsrechnung eignen sich zur quantitativen Beurteilung von Investitionen und Entscheidungen.⁴⁸⁴ Qualitative Aspekte, wie bspw. die Präferenzen von Entscheidungsträgern, finden bei den qualitativen Entscheidungsmodellen Berücksichtigung.⁴⁸⁵ Sowohl quantitative als auch qualitative Aspekte können bei der Beurteilung von Investitionen einbezogen werden.⁴⁸⁶ Hierfür eignet sich die Anwendung der Nutzwertanalyse.⁴⁸⁷ Die Nutzwertanalyse ermittelt auf Basis subjektiver Zielgewichtungen und Teilnutzenbestimmungen einen Gesamtnutzwert.⁴⁸⁸ Ziel der Nutzwertanalyse ist der Vergleich von Investitionsalternativen.⁴⁸⁹ Die Vereinheitlichung und Vergleichbarkeit der einzelnen Kriterien der Investitionsalternativen erfolgt durch Vergabe von Nutzwerten.⁴⁹⁰ Die Investitionsalternative mit dem höchsten Nutzwert wird der Investitionsalternative mit geringerem Nutzwert vorgezogen. Es können bei der Nutzwertanalyse allein qualitative Aspekte berücksichtigt werden und quantitativen Zielgrößen gegenübergestellt werden.⁴⁹¹ Alternativ können quantitative Ergebnisse auch in die Nutzenbewertung einbezogen werden.⁴⁹² Der Vorteil der Nutzwertanalyse liegt darin, dass neben quantitativen Größen zusätzlich qualitative Aspekte in die Entscheidungsvorbereitung einfließen. Quantitative Daten lassen sich allerdings einfacher erheben, als qualitative Daten.⁴⁹³ Die Nutzwerte basieren auf Schätzungen, weshalb entsprechende Ungenauigkeiten in die Bewertung der Investitionsalternativen einfließen können.⁴⁹⁴

⁴⁸⁴ Vgl. Kapitel 4.2.1, Kapitel 4.2.2; Diese Modelle und Methoden basieren auf monetären Größen, Zahlenwerten und Kennzahlen.

⁴⁸⁵ Vgl. Hennink, Hutter, Bailey / Qualitative Research / 9-10

⁴⁸⁶ Vgl. Zantow, Dinauer / Finanzwirtschaft / 477

⁴⁸⁷ Vgl. Zangemeister / Nutzwertanalyse / 45; Hoffmeister / Investitionsrechnung / 296

⁴⁸⁸ Vgl. Zantow, Dinauer / Finanzwirtschaft / 478

⁴⁸⁹ Vgl. Jung / Controlling / 134

⁴⁹⁰ Als Nutzwerte dienen Punktwerte, die als Kriterien-Wert der jeweiligen Alternative vom Entscheider zugeordnet werden. Vgl. hierzu Mehlan / Controlling / 55.

⁴⁹¹ Hierbei handelt es sich um den Nutzwert im engeren Sinne. Vgl. hierzu Dreger / Kundenzufriedenheit / 153.

⁴⁹² Hierbei handelt es sich um den Nutzwert im weiteren Sinne. Vgl. hierzu Zantow, Dinauer / Finanzwirtschaft / 478.

⁴⁹³ Vgl. Riesenhuber / Empirische Forschung / 7

⁴⁹⁴ Vgl. Wünsche / Finanzwirtschaft / 18; Hoffmeister / Investitionsrechnung / 307

4.3 Wirtschaftlichkeitsanalyse im ILM

Ergebnisse empirischer Untersuchungen haben gezeigt, dass zwischen zwanzig und siebenzig Prozent der von Unternehmen getätigten Investitionen in die IT keinen Nutzen generieren.⁴⁹⁵ Wesentliche Gründe hierfür sind bspw. der Mangel an qualifiziertem Personal,⁴⁹⁶ die fehlende Abstimmung zwischen den Projektbeteiligten⁴⁹⁷, das Fehlen von definierten Anforderungen und Zielstellungen⁴⁹⁸ sowie die ungenügende Analyse der relevanten Kosten- und Nutzengrößen.⁴⁹⁹ Zur Beurteilung der Vorteilhaftigkeit von ILM ist die Durchführung einer Wirtschaftlichkeitsanalyse⁵⁰⁰ unerlässlich.⁵⁰¹ Im Folgenden wird ein Bezug zwischen ILM und der Wirtschaftlichkeitsanalyse hergestellt. Wenn der Nutzen größer als die Kosten ist, dann ist ILM wirtschaftlich.⁵⁰² Bei den Kosten wird unterschieden zwischen einmaligen und wiederkehrenden Kosten.⁵⁰³ Beispiele für einmalige Kosten sind IT-Anschaffungskosten, Kosten für die Ausbildung von Mitarbeitern oder Kosten für die Einrichtung von IT-Systemen.⁵⁰⁴ Betriebs- und Wartungskosten wie Miet-, Versicherungs- oder Datenübertragungskosten sind wiederkehrende Kosten.⁵⁰⁵ Die Ermittlung des Nutzens gestaltet sich schwieriger, als die Ermittlung der Kosten.⁵⁰⁶ Die Erträge und Leistungen resultieren in der Regel nicht ausschließlich aus monetär messbarem Nutzen, sondern zusätzlich aus dem indirekt messbaren Nutzen und

⁴⁹⁵ Vgl. ITGI / Val IT Framework / 7; Cook / Failing Project / 1-6

⁴⁹⁶ Vgl. Krcmar / Informationsmanagement / 145

⁴⁹⁷ Vgl. Filz / IT-Kompetenz / 51-52

⁴⁹⁸ Vgl. Beck et al. / Requirements Engineering / 19-20

⁴⁹⁹ Vgl. hierzu ITGI / Val IT Framework / 9-10; Rüter et al. / IT-Governance / 168-171; Nehfort / IT-Lösungen / 495. Ziel von ILM ist die kostengünstige Verwaltung von Daten (Vgl. Kapitel 1.1). Dennoch wurde die Analyse der Wirtschaftlichkeit der automatisierten Verwaltung unstrukturierter Daten im ILM in wissenschaftlichen Publikationen bisher vernachlässigt. Vgl. hierzu Kapitel 1.1 und Kapitel 5.1.3.1.

⁵⁰⁰ Vgl. zur Durchführung einer Wirtschaftlichkeitsanalyse Killich, Maicher, Pfeiffer / IT-Value / 61-68; ITGI / Val IT Framework / 45; Bundesministerium des Inneren / WiBe / 8-10, 70-84; Gadatsch, Mayer / IT-Controlling / 291-300.

⁵⁰¹ Vgl. hierzu Feyhl / Controlling von Softwareprojekten / 109-110; Krüger / Marktfähige Produkte / 151-152; Wieczorrek, Mertens / Management IT-Projekte / 225-227; Marx Gómez, Junker, Odebrecht / IT-Controlling / 292. Schwerpunkt der vorliegenden Arbeit ist die Wirtschaftlichkeitsanalyse der automatisierten Klassifizierung unstrukturierter Daten im ILM. Vgl. hierzu Kapitel 1.1.

⁵⁰² Vgl. Kapitel 4.1 zur Definition der Begriffe Kosten und Nutzen und Kapitel 5.3 zur Durchführung bzw. Demonstration der Wirtschaftlichkeitsanalyse.

⁵⁰³ Vgl. Zarnekow, Brenner / Kosten IT / 336-339

⁵⁰⁴ Vgl. Viereck, Sonderhüsen / Informationstechnik in der Praxis / 408; Zarnekow / IT-Dienstleistungen / 118-119

⁵⁰⁵ Vgl. Amberg, Wiener / IT-Offshoring / 64-67

⁵⁰⁶ Vgl. Biethahn, Mucksch, Ruf / Informationsmanagement / 360-361

dem nicht messbaren Nutzen.⁵⁰⁷ Direkt messbarer Nutzen⁵⁰⁸ entsteht durch Kostensenkung, bspw. wenn ein altes durch ein neues Anwendungssystem ersetzt wird.⁵⁰⁹ Führen bspw. IT-Investitionen zu einer Produktivitätssteigerung oder steigern sie die Effizienz von Unternehmensprozessen, so entsteht ein indirekt messbarer Nutzen.⁵¹⁰ Profitieren verschiedene Unternehmensbereiche von IT-Investitionen, können bspw. unternehmensweite Kommunikationswege verbessert werden, was zu qualitativ höherwertigen Entscheidungen und zu monetär nicht messbarem Nutzen führen kann.⁵¹¹ Ziel von ILM ist es, Informationen zu klassifizieren und kostengünstig zu verwalten.⁵¹² Wenn Informationen, die einen niedrigen Nutzungsgrad⁵¹³ aufweisen, von kostenintensiven operativen Speichermedien auf kostengünstige Speichermedien verlagert werden,⁵¹⁴ entsteht durch Kostensenkung ein direkt messbarer Nutzen. Durch die Entlastung der operativen Systeme kann zusätzlich ein indirekter Nutzen entstehen, da bspw. Unternehmensprozesse schneller abgearbeitet werden können und die Produktivität erhöht wird.⁵¹⁵ Die Analyse des Kosten- und Nutzenverhältnisses, das bei der Implementierung und Anwendung von ILM sowie der Verwaltung unstrukturierter Daten entsteht, wird in der Literatur bislang vernachlässigt.⁵¹⁶ Die Verfasser konzentrieren sich auf die Entwicklung von Klassifizierungsfunktionen, um den Informationen entsprechend ihres Nutzungsgrads einen Klassifizierungswert zuzuweisen.⁵¹⁷ Welche Kosten allerdings bei der Implementierung dieser Konzepte entstehen und welcher Nutzen daraus entsteht, wird

⁵⁰⁷ Der Nutzen von IT-Investitionen wirkt sich auf sämtliche Unternehmensbereiche aus. Neben der Kostensenkung und der Erhöhung der Produktivität lassen sich strategische Wettbewerbsvorteile erzielen, wenn die Unternehmensziele und die Ziele der IT aufeinander abgestimmt sind. Vgl. hierzu ISACA / Business Framework / 27-30.

⁵⁰⁸ Hierfür eignet sich bspw. die Anwendung der Kostenvergleichsrechnung. Vgl. hierzu Kapitel 4.2.1.1.

⁵⁰⁹ Vgl. Wiczorrek, Mertens / Management IT-Projekte / 228

⁵¹⁰ Vgl. Brugger / IT Business Case / 92; Bräu, Dannert / Effizienzsteigerung / 260

⁵¹¹ Vgl. Wiczorrek, Mertens / Management IT-Projekte / 228; Marx Gómez, Junker, Odebrecht / IT-Controlling / 190-191

⁵¹² Vgl. Maier, Hädrich, Peinl / Knowledge Infrastructures / 254-257; Thome, Sollbach / Modelle des ILM / 22-30

⁵¹³ Vgl. Kapitel 2.4.1

⁵¹⁴ Vgl. Kapitel 2.4.2

⁵¹⁵ Vgl. Kapitel 2.4.2

⁵¹⁶ Vgl. Kapitel 1.1 und Kapitel 5.1.3.1.2

⁵¹⁷ Vgl. Chen / Information valuation / 145-146; Verma et al. / Very Large File Systems / 166-167; Turczyk / ILM / 117-118

nicht betrachtet.⁵¹⁸ Die Verfasser der Konzepte machen wenige bzw. keine Aussagen darüber, ob die ILM-Konzepte bei großen Datenbeständen praktikabel einsetzbar sind.⁵¹⁹

ILM eignet sich ebenfalls für die Verwaltung strukturierter Daten. Im Vordergrund der Betrachtungen stehen technische und funktionale Möglichkeiten zur Entlastung⁵²⁰ der operativen Datenbanken durch die Verlagerung der Daten in Nearline-Storage-Systeme oder Archivsysteme.⁵²¹ Beispiele für Datenbanken mit strukturierten Daten sind ERP-Systeme oder Data-Warehouse-Systeme. Wie die Kriterien zur Unterstützung der automatisierten Klassifizierung strukturierter Daten ermittelt werden können, wird ebenfalls in der Literatur diskutiert.⁵²² Der Aufwand zur Bestimmung der Klassifizierungskriterien ist im Vergleich zu unstrukturierten Daten höher. Bei strukturierten Datensätzen ist die Ermittlung des Nutzungsgrads aufgrund der in den Datenbanken vorkommenden starken Interdependenzen wesentlich schwieriger.⁵²³ Nur wenige Arbeiten beschäftigen sich mit der Wirtschaftlichkeitsanalyse strukturierter Daten im ILM.⁵²⁴ Da der Großteil der Unternehmensdaten unstrukturiert ist,⁵²⁵ ist das Potential zur Senkung der Speicher- und Administrationskosten bei der Verwaltung unstrukturierter Daten entsprechend hö-

⁵¹⁸ Vgl. hierzu Kapitel 5.1.3.1.2. Die Verfasser der ILM-Konzepte weisen darauf hin, dass durch den Einsatz von ILM operative Speicher entlastet werden können. Allerdings wird hierbei nicht untersucht, ob und welcher monetärer oder nicht monetärer Nutzen dadurch entsteht. Speicherhersteller und Beratungshäuser liefern ebenfalls keine konkreten Informationen zum Nutzen von ILM. Das liegt vor allem darin begründet, dass vollständige ILM-Implementierungen in der Praxis bislang noch fehlen. Vgl. hierzu SNIA / ILM Model / 5; Thome, Sollbach / Modelle des ILM / 150; Kaiser, Smolnik, Riempp / ILM-Framework / 483-484.

⁵¹⁹ Vgl. Kapitel 5.1.3.1.2

⁵²⁰ Mit dem wesentlichen Ziel, die Leistungsfähigkeit der operativen Systeme aufrecht zu erhalten bzw. die Anforderungen an Antwortzeiten und die Verfügbarkeit der Daten zu erfüllen.

⁵²¹ Vgl. Inmon / Data Warehouse / 31; Steinbinder, Bedau, Löw / Datenbanktechnologie für das ILM / 101-114; Jeong et al. / Effective Storage / 30-46; Liu, Wang, Quan / ILM Architecture / 166-168; Badam, Kumari, Srinivasan / ILM healthcare / 79-90

⁵²² Ein Ansatz zur Ermittlung der Zugriffe auf einzelne Datensätze im SAP Business Information Warehouse wird in Kosler, Matthesius, Stelzer / Klassifizierung von Informationen / 129-145 vorgestellt.

⁵²³ Vgl. Ballou, Smith, Wang / Quality Estimation Database / 639-650; Chaudhuri, Weikum / Database System Architecture / 1-10; Kaiser, Smolnik, Riempp / ILM-Framework / 484; Vassiliadis et al. / Warehouse Process Management / 205-236; Kosler, Matthesius, Stelzer / Klassifizierung von Informationen / 129-145

⁵²⁴ Vgl. Kapitel 1.3.2. Der Nutzen von ILM konnte hierbei für Data-Warehouse-Systeme im Rahmen von prototypischen Implementierungen, für Unternehmen der Energiebranche, in Form von Kostensenkungen aufgezeigt werden. Vgl. hierzu Matthesius / ILM-Konzept Energiekonzern / 38-65; Oßmann / Daten ILM / 23-24; Heinrich, Stelzer / Informationsmanagement / 525-536.

⁵²⁵ Vgl. Kapitel 1.1

her als bei der Verwaltung strukturierter Daten im ILM. Ziel der vorliegenden Arbeit ist die Analyse der Wirtschaftlichkeit der automatisierten Verwaltung unstrukturierter Daten im ILM. Es wird für unterschiedliche Lebenszyklen und Datenmengen untersucht, welche Kosten bei der Implementierung von ILM anfallen und welcher Nutzen dabei entsteht.⁵²⁶

5 Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM

5.1 Konstruktion eines Vorgehensmodells für das ILM

5.1.1 Vorstellung von ILM-Vorgehensmodellen

In diesem Kapitel werden bereits existierende ILM-Vorgehensmodelle vorgestellt. Aufbauend auf den Vorgehensmodellen wird in Kapitel 5.1.2 ein ILM-Vorgehensmodell konstruiert,⁵²⁷ das die wesentlichen Teilaufgaben des ILM beschreibt und neben der Verwaltung unternehmensinterner Speichermedien zusätzlich das Konzept des Cloud Computings berücksichtigt.⁵²⁸ Teile des ILM-Vorgehensmodells bilden die konzeptionelle Basis für die in dieser Arbeit durchgeführte Wirtschaftlichkeitsanalyse.⁵²⁹ Es werden Anforderungen an die Wirtschaftlichkeitsanalyse im ILM definiert⁵³⁰ und Klassifizierungs-, Verlagerungs- und Kostenfunktionen implementiert⁵³¹. Anschließend wird eine Kostenvergleichsrechnung⁵³² durchgeführt. Das Kosten- und Nutzenverhältnis für unterschiedliche Lebenszyklen bzw. Nutzungsgrade wird ermittelt.⁵³³ Die aus der Durchführung bzw. Demonstration der Wirtschaftlichkeitsanalyse resultierenden Ergebnisse werden im Hinblick auf die zuvor definierten Anforderungen evaluiert.⁵³⁴

⁵²⁶ Vgl. Kapitel 5.2.1 und Kapitel 5.3

⁵²⁷ Das Vorgehensmodell enthält die Teilphasen „Analyse der IS-Architektur“, „Beschreibung relevanter Datenbestände“, „Klassifizierung der Daten“ und „Verlagerung der Daten“.

⁵²⁸ Vgl. Kapitel 1.2; Kapitel 1.3; Im Rahmen von ILM stellt das Cloud Computing eine weitere Möglichkeit zur kostengünstigen Verwaltung von Informationen dar. Vgl. hierzu Kapitel 1.1 und Kapitel 3.

⁵²⁹ Vgl. Kapitel 1.3; Kapitel 5.2; Kapitel 5.3

⁵³⁰ Vgl. Kapitel 5.2.1

⁵³¹ Vgl. Kapitel 5.2.3.1, Kapitel 5.2.3.2, Kapitel 5.2.3.3

⁵³² Vgl. Kapitel 4.2.1.1

⁵³³ Vgl. Kapitel 5.2.2, Kapitel 5.3

⁵³⁴ Vgl. Kapitel 5.4

Durch die Analyse relevanter Literatur aus Zeitschriften und Konferenzen wurden die Vorgehensmodelle von StorageTek und EMC identifiziert.⁵³⁵ Weitere zwei Vorgehensmodelle, publiziert durch die SNIA und den BITKOM, wurden mit Hilfe von Suchmaschinen-Diensten ermittelt.⁵³⁶ Diese ILM-Vorgehensmodelle werden im Folgenden vorgestellt. Das Ziel des ILM-Vorgehensmodells des BITKOM⁵³⁷ ist die regelmäßige Klassifikation⁵³⁸ der Informationen, um eine möglichst kostengünstige Verwaltung der Informationen zu gewährleisten. Berücksichtigt werden der Wert und das Alter der Informationen, die Speicherkosten sowie gesetzliche Anforderungen. Das Vorgehensmodell ist in Abbildung 5-1 abgebildet. Es besteht aus den Phasen „Erstplatzierung und Datensicherung“, „Verdrängung“, „Archivierung“ sowie „Löschen“.

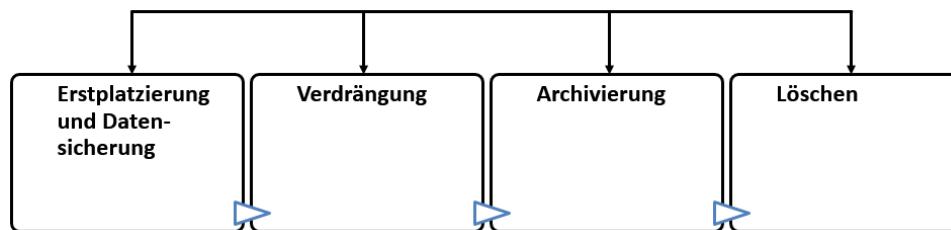


Abb. 5-1: ILM-Vorgehensmodell BITKOM⁵⁴⁰

In der Phase „Erstplatzierung und Datensicherung“ werden die Informationen entsprechend ihres Werts klassifiziert und auf den dafür vorgesehenen Speichermedien abgelegt. Weiterhin erfolgt eine regelmäßige Sicherung der Informationen, um den dauerhaften Zugriff zu gewährleisten. Wenn sich der Wert der Informationen im Lebenszyklus ändert, werden die Informationen mit Hilfe der „Verdrängung“ auf eine dafür vorgesehene Ebene der Speicherhierarchie⁵⁴¹ verlagert. Ist eine langfristige Aufbewahrung der

⁵³⁵ Vgl. hierzu Gießelbach / ILM / 7-24, Reiner et al. / ILM EMC / 804 und die Literaturanalyse in Kapitel 1.3.2.

⁵³⁶ Vgl. Kapitel 1.3.2

⁵³⁷ Vgl. BITKOM / Leitfaden ILM / 7-15

⁵³⁸ Die Daten müssen in regelmäßigen Abständen manuell bewertet werden. Allerdings liefert der BITKOM hinsichtlich einer Zeitabschätzung keine Hinweise, was „regelmäßig“ genau bedeutet.

⁵³⁹ Es wird auf Basis der Datenklassifizierung festgelegt, in welcher Speicherebene eine Information gespeichert werden soll. Vgl. hierzu BITKOM / Leitfaden ILM / 10.

⁵⁴⁰ In Anlehnung an BITKOM / Leitfaden ILM / 12

⁵⁴¹ Vgl. Kapitel 2.4.2, Abbildung 2-2

Informationen erforderlich, findet die „Archivierung“⁵⁴² statt. Am Ende des Lebenszyklus erfolgt das „Löschen“ der Informationen, um den Speicherplatz freizugeben. Die einzelnen Phasen können mehrfach durchlaufen werden. Bereits gespeicherte oder archivierte Informationen können erneut klassifiziert werden und einer entsprechenden Speicherebene zugeordnet oder gelöscht werden, wenn sich der Wert der Informationen ändert. Der BITKOM macht keine Angaben dazu, wie der Wert der Informationen bestimmt werden kann. Es wird lediglich darauf hingewiesen, dass die Festlegung des Werts als unternehmerische Vorgabe erfolgen muss. Der Fokus des Vorgehensmodells liegt auf dem Lebenszyklus von Informationen und den jeweils relevanten Speichermedien. Die wesentlichen Teilaufgaben des ILM werden nicht beschrieben. Wie bei der Umsetzung eines ILM-Konzepts vorgegangen werden sollte, wird nicht näher erläutert.

Das in Abbildung 5-2 dargestellte ILM-Vorgehensmodell des Speicherherstellers StorageTek, das ebenfalls die kostengünstige Verwaltung von Informationen ermöglichen soll, besteht aus den fünf Phasen „Verteilen“, „Bereitstellen“, „Schützen“, „Konservieren“ und „Archivieren“.⁵⁴³ In den Phasen „Verteilen“ und „Bereitstellen“ erfolgt, nachdem die Informationen erstellt wurden, die Verteilung auf die verschiedenen operativen Systeme. Anschließend werden die Informationen in der Phase „Schützen“ mit Zugriffsberechtigungen versehen, damit ein Zugriff nur durch autorisierte Personen möglich ist. Sinkt der Wert im Zeitverlauf, erfolgt in Phase vier das „Konservieren“ der Informationen, also die Verlagerung auf kostengünstige Speichermedien. Die „Archivierung“ sorgt schließlich für die langfristige Aufbewahrung der Informationen in der Offline-Ebene. Inwiefern Informationen gelöscht werden sollen, wird im Vorgehensmodell nicht erwähnt.

⁵⁴² Vgl. Kapitel 2.4.2

⁵⁴³ Vgl. Gießelbach / ILM / 7-24

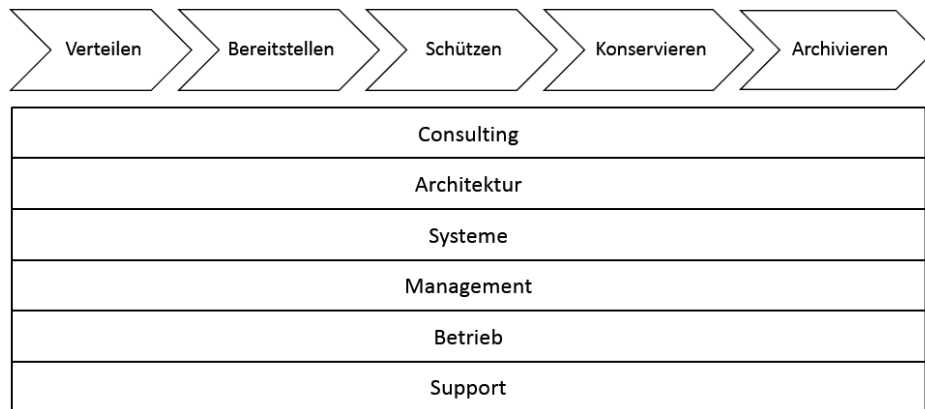


Abb. 5-2: ILM-Vorgehensmodell StorageTek⁵⁴⁴

Das Vorgehensmodell beschreibt nicht ausschließlich die einzelnen Phasen des Lebenszyklus von Informationen und die jeweils relevanten Speicherebenen. Es wird zusätzlich aufgezeigt, dass die Bereiche bzw. Teilphasen „Consulting“, „Architektur“, „Systeme“, „Management“, „Betrieb“ und „Support“ für ILM bzw. die einzelnen Lebenszyklusphasen von Informationen von Relevanz sind. Die Analyse und Beratung zu Zielen und Nutzenpotentialen, idealerweise durchzuführen von externen Beratern, erfolgt in der Teilphase „Consulting“. Der Wert der Informationen wird festgelegt und relevante Datenbestände werden untersucht. Wie der Wert der Informationen bestimmt werden soll, wird nicht beschrieben. Durch die Analyse der „Architektur“⁵⁴⁵ und der in einem Unternehmen vorhandenen „Systeme“ wird untersucht, welche Hardware und welche unternehmensinternen Speichermedien für die Verwaltung der Daten in der jeweiligen Phase des Lebenszyklus zum Einsatz kommen sollen. Wie die Verwaltung und der Betrieb der Speichermedien erfolgen sollen, wird in den Teilphasen „Management“ und „Betrieb“ festgelegt. Die Teilphase „Support“ regelt Aspekte der Wartung von Speichermedien sowie die Schulung von Mitarbeitern zur Administration der Speichermedien und zum Konzept des ILM. Das ILM-Vorgehensmodell fokussiert stark auf den Vertrieb von Speicherhardware und Beratungsdienstleistung. Dennoch wird aufgezeigt, welche betrieblichen Bereiche⁵⁴⁶ durch den Lebenszyklus von Informationen tangiert werden. Auch die SNIA⁵⁴⁷ und der Speicherhersteller EMC⁵⁴⁸ weisen in ihren ILM-

⁵⁴⁴ In Anlehnung an Gießelbach / ILM / 12

⁵⁴⁵ Gemeint ist die Hardware-Architektur, bspw. welche Rechner und Server zur Verwaltung der Informationen genutzt werden.

⁵⁴⁶ Vgl. Abbildung 5-2

⁵⁴⁷ Vgl. SNIA / ILM Definition / 4-10

Vorgehensmodellen darauf hin, dass die Informationen nach ihrem Wert klassifiziert werden müssen, um diese anschließend automatisiert mit Hilfe der ILM-Software auf dem jeweils kostengünstigsten Speichermedium bereitstellen zu können. Abbildung 5-3 enthält das Vorgehensmodell der SNIA.

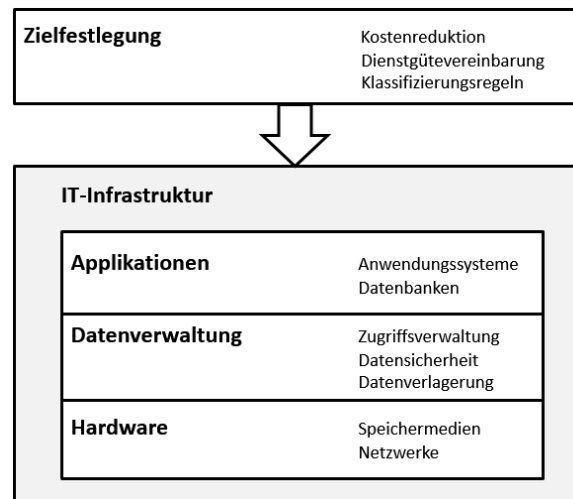


Abb. 5-3: ILM-Vorgehensmodell SNIA⁵⁴⁹

Ähnlich wie bei StorageTek erfolgt im Vorgehensmodell der SNIA zu Beginn die „Zielfestlegung“ für das ILM. Von zentraler Bedeutung ist die Reduktion von Speicher- und Administrationskosten. Durch die Festlegung von Klassifizierungsregeln werden geschäftsrelevante Informationen von Informationen mit geringerem Wert getrennt und mit jeweils geeigneten unternehmensinternen Speichermedien verwaltet. Die Klassifizierung und anschließende Verlagerung der Informationen muss im Einklang mit getroffenen Dienstgütevereinbarungen erfolgen, sodass Vorgaben zu bspw. Verfügbarkeit der Informationen und Antwortzeiten beim Abruf der Informationen eingehalten werden. Nachdem die Ziele festgelegt wurden, erfolgt die Untersuchung und Anpassung der „IT-Infrastruktur“. Relevante Anwendungssysteme und Datenbanken werden analysiert. Dabei wird geprüft, ob bestehende Zugriffsrechte auf die Informationen auch nach der Verlagerung auf unterschiedliche Speichermedien, ggf. über verschiedene Netzwerke, erhalten bleiben. Somit soll die Datensicherheit gewährleistet werden. Das Vorgehensmodell der SNIA leitet die Klassifizierungsregeln aus Zielvorgaben und Dienstgütevereinbarungen ab. Somit kann bspw. sichergestellt werden, dass geschäftsrelevante In-

⁵⁴⁸ Vgl. James / Accelerating ILM / 3-5; Reiner et al. / ILM EMC / 804

⁵⁴⁹ In Anlehnung an SNIA / ILM Definition / 7

formationen für zukünftige Projekte zur Verfügung stehen und nicht aufgrund der Klassifizierungsregeln gelöscht werden. Im Rahmen des Modells wird nicht erörtert, wie der Geschäftswert von Informationen ermittelt werden kann. Zentrales Element des Vorgehensmodells ist die Verwaltung der IT-Infrastruktur und der eingesetzten Speichermedien. EMC verfolgt einen ähnlichen Ansatz zur Klassifizierung und Verwaltung von Informationen im ILM. Zunächst erfolgt die „Klassifizierung der Anwendungssysteme und Daten“. Die Informationen werden anhand ihres Werts klassifiziert. Die Implementierung der Klassifizierungsregeln, die später für die automatisierte Klassifizierung der Informationen sorgen, erfolgt in der Phase „Implementierung der Klassifizierungsregeln“. In den Phasen „Verwaltung der Speicherumgebung“ und „Einteilung der Speichermedien in Speicherklassen“ werden relevante Speichermedien zur Verwaltung der Informationen identifiziert und die Speichermedien in die zuvor gebildeten Klassen eingeteilt. Das Vorgehensmodell von EMC ist in Abbildung 5-4 dargestellt.

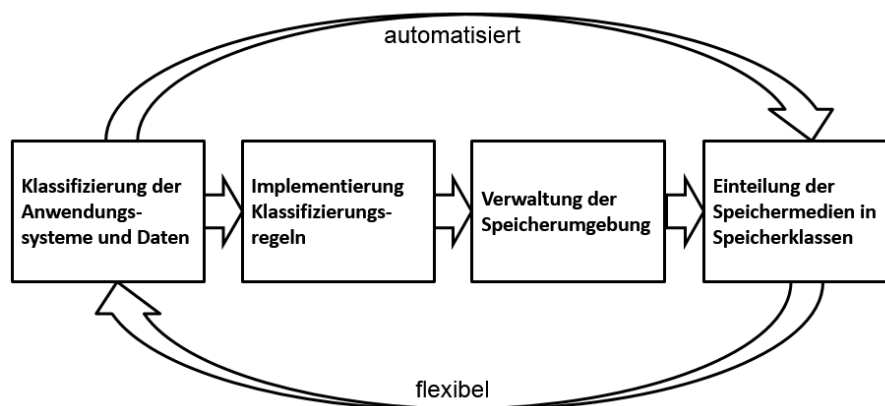


Abb. 5-4: ILM-Vorgehensmodell EMC⁵⁵⁰

Der Schwerpunkt des EMC-Vorgehensmodells liegt in der „automatisierten“ Verwaltung der IT-Infrastruktur und der unternehmensinternen Speichermedien. Durch eine „flexible“ Verwaltung wird die Anpassung der Speichermedien und Speicherklassen an Änderungen in den Klassen und Klassifizierungsregeln ermöglicht.⁵⁵¹

⁵⁵⁰ In Anlehnung an James / Accelerating ILM / 2

⁵⁵¹ Bspw. wenn neue Klassen hinzukommen.

Die aufgeführten Vorgehensmodelle für ILM haben noch nicht den Stand erreicht, sich in der Praxis durchzusetzen,⁵⁵² da umfangreiche und vollständige Implementierungen bislang fehlen.⁵⁵³ Die Beschreibung der Teilaufgaben des ILM bezieht sich hauptsächlich auf die Verwaltung der eingesetzten Speichermedien. Weiterhin besteht innerhalb der Vorgehensmodelle Unklarheit darüber, was unter dem „Wert von Informationen“ zu verstehen ist bzw. worin sich dieser äußert.⁵⁵⁴ Keines der ILM-Vorgehensmodelle berücksichtigt Cloud Computing als Möglichkeit zur Verwaltung von Informationen. Bislang existieren wenige wissenschaftliche Untersuchungen dazu, wie das Cloud Computing im ILM eingesetzt werden kann, um Informationen zu verwalten.⁵⁵⁵ Die Nutzung des Cloud Computings, bspw. als Dienst über das Internet, kann jedoch kostengünstiger sein, als die Verwaltung der Informationen auf unternehmenseigenen Speichermedien.⁵⁵⁶ Aus diesem Grund sollte das ILM erweitert werden, indem das Cloud Computing, als zusätzliche Möglichkeit zur Verwaltung von Informationen, mit einbezogen wird.

5.1.2 Konstruktion eines ILM-Vorgehensmodells

Aufbauend auf den in Kapitel 5.1.1 vorgestellten Vorgehensmodellen wird in diesem Kapitel ein ILM-Vorgehensmodell konstruiert.⁵⁵⁷ Eine weitere Grundlage für das zu konstruierende Vorgehensmodell bildet die Literaturanalyse zum Information Lifecycle Management und Cloud Computing.⁵⁵⁸ Welche Merkmale aus den bereits existierenden Vorgehensmodellen in das zu konstruierende Modell übernommen werden und welche Merkmale zusätzlich definiert werden, wird im Folgenden erörtert. Die bereits vorgestellten Vorgehensmodelle weisen die folgenden Merkmale auf, die auch dem zu konstruierenden Vorgehensmodell zugrunde gelegt werden:

⁵⁵² Vgl. Thome, Sollbach / Modelle des ILM / 150

⁵⁵³ Vgl. Kaiser, Smolnik, Riempp / ILM-Framework / 483-484

⁵⁵⁴ Vgl. Kapitel 2.4.1

⁵⁵⁵ Vgl. Du / Optimal Storage / 730-746; Sotomayor et al. / Private and Hybrid Clouds / 14-21; Jo et al. / Virtual Disk / 375-377; Hill et al. / Guide Cloud Computing / 250-251

⁵⁵⁶ Wenn Informationen im Cloud Computing als Dienst über das Internet verwaltet werden, können die Kosten für die Administration unternehmenseigener Speichermedien weiter gesenkt werden. Vgl. hierzu Schadler / Email in the Cloud / 3-5.

⁵⁵⁷ Vgl. Kapitel 1.3.1

⁵⁵⁸ Vgl. Kapitel 1.3.2; Vgl. zur Erstellung von Modellen auf Basis einer Literaturanalyse Goeken / Ansätze der Referenzmodellierung / 7.

- Einteilung des Vorgehens in Phasen: Ein Vorgehensmodell beschreibt das systematische Vorgehen in gleichartigen Projekten,⁵⁵⁹ um bestimmte Probleme einer Klasse wiederholbar zu lösen⁵⁶⁰.⁵⁶¹ Mit Hilfe von Phasen wird in einem Vorgehensmodell der Ablauf von bspw. Entwicklungsvorhaben und Projekten beschrieben.⁵⁶² Die Phasen fassen bestimmte Tätigkeiten bzw. Teilschritte eines Projekts zusammen.⁵⁶³ Bei allen in Kapitel 5.1.1 vorgestellten Vorgehensmodellen zum ILM erfolgt eine Einteilung des Vorgehens in Phasen.⁵⁶⁴
- Iteration der Phasen: Die einzelnen Phasen des Vorgehensmodells können mehrfach durchlaufen werden. Der BITKOM und EMC weisen in ihren Vorgehensmodellen darauf hin, dass bereits gespeicherte oder archivierte Informationen erneut klassifiziert und einer entsprechenden Speicherebene zugeordnet oder gelöscht werden können, wenn sich der Wert der Informationen ändert.
- Untersuchung der Anwendungssysteme und Datenbestände: Die in einem Unternehmen vorhandenen Anwendungssysteme und Datenbestände müssen ermittelt und untersucht werden. Auf dieser Basis werden die Informationen klassifiziert.⁵⁶⁵
- Einteilung der Informationen und Speichermedien in Klassen: Informationen müssen anhand ihres Werts klassifiziert werden. Die Klassifizierung sollte automatisiert erfolgen. Entsprechend des Werts müssen die Informationen auf jeweils sinnvollen Speichermedien verwaltet werden.⁵⁶⁶

⁵⁵⁹ Vgl. Brandt-Pook, Kollmeier / Softwareentwicklung / 3-4

⁵⁶⁰ Vgl. Broy, Kuhrmann / Projektorganisation / 86

⁵⁶¹ Es beschreibt den Prozess bzw. die Vorgehensweise zur Lösung eines Problems. Vgl. hierzu Ferstl, Sinz / Wirtschaftsinformatik / 468-469; Heinrich, Stelzer / Informationsmanagement / 403-405.

⁵⁶² Vgl. Becker et al. / Referenzvorgehensmodell / 153; Allweyer / Geschäftsprozessmanagement / 97. Vgl. zu den Vorgehensmodellen der Softwareentwicklung Badertscher, Gubelmann, Scheuring / Informationssysteme / 165-169; Heinrich / Systemanalyse / 63-81

⁵⁶³ Vgl. Becker, Meise / Strategie und Ordnungsrahmen / 115-116; Becker et al. / Referenzmodellierung / 34-36

⁵⁶⁴ Vgl. hierzu Abbildung 5-1 bis Abbildung 5-4.

⁵⁶⁵ Vgl. hierzu die Vorgehensmodelle von StorageTek, SNIA und EMC.

⁵⁶⁶ Vgl. Abbildung 5-1 bis Abbildung 5-4

- Berücksichtigung von Anwendern und Administratoren: Das Wissen von Anwendern und Administratoren, bspw. zur aktuellen und zukünftigen Nutzung der Informationen, ist zu berücksichtigen. Somit können Vorgaben zur Verfügbarkeit der Informationen und zu Antwortzeiten beim Abruf der Informationen eingehalten werden.

In keinem der beschriebenen Vorgehensmodelle wird erläutert, wie bspw. der Wert von Informationen ermittelt werden kann oder wie die Informationen klassifiziert werden. Wesentlicher Bestandteil von ILM ist jedoch die Klassifizierung von Informationen.⁵⁶⁷ Schwerpunkt der Modelle bildet die Verwaltung von Speichermedien während des Informationslebenszyklus. Um eine Wirtschaftlichkeitsanalyse für das ILM durchführen zu können, müssen bspw. die wesentlichen Teilaufgaben des ILM beschrieben werden, um die jeweiligen Kosten- und Nutzenaspekte ableiten zu können.⁵⁶⁸ Die folgenden zusätzlichen Merkmale werden im zu konstruierenden Vorgehensmodell berücksichtigt:

- Beschreibung der wesentlichen Teilaufgaben des ILM: Alle vorgestellten ILM-Vorgehensmodelle haben das Ziel, die Kosten für die Verwaltung von Speichermedien zu senken. Die wesentlichen Teilaufgaben des ILM müssen beschrieben werden.⁵⁶⁹ Daraus können Anschaffungskosten, Betriebskosten und sonstige Kosten abgeleitet werden. Die in Kapitel 5.1.1 vorgestellten Vorgehensmodelle beschreiben in den jeweiligen ILM-Phasen entweder die Notwendigkeit zur Klassifizierung von Informationen im Informationslebenszyklus⁵⁷⁰ oder die Bestandteile der IT-Infrastruktur zur Verwaltung von Speichermedien⁵⁷¹. Neben diesen Aspekten sollten zusätzlich das allgemeine Vorgehen und die entsprechenden Teilschritte zur Umsetzung von ILM erörtert werden.⁵⁷²

⁵⁶⁷ Vgl. Abd-El-Malek et al. / self-storage / 55-62; Bhagwan et al. / Management of Data Storage /; Chen / Information valuation / 135-146; Mesnier et al. / File classification / 44-51; Thome, Sollbach / Modelle des ILM / 22-30; Yates-Mercer, Bawden / Managing the paradox / 19-29

⁵⁶⁸ Vgl. Kapitel 4.3

⁵⁶⁹ Es wird untersucht, inwiefern ILM überhaupt zu einer Verbesserung der Wirtschaftlichkeit bei der Verwaltung von unstrukturierten Daten, mit verschiedenen Lebenszyklen, beiträgt. Vgl. zur Wirtschaftlichkeit und Potentialanalyse bei der Erstellung und Verwendung von Modellen Becker, Probandt, Vering / Grundsätze ordnungsgemäßer Modellierung / 34-35; vom Brocke, Buddendick / Konstruktionstechniken Referenzmodellierung / 23-25; Becker et al. / Referenzmodellierung / 26.

⁵⁷⁰ Vgl. Abbildung 5-1 und 5-2

⁵⁷¹ Vgl. Abbildung 5-3 und 5-4

- Beschreibung der Vorgehensweise zur Klassifizierung der Informationen: Wie und auf Basis welcher Kriterien der Wert der Informationen ermittelt werden kann und wie die Informationen klassifiziert werden können, sollte Bestandteil eines ILM-Vorgehensmodells sein.
- Berücksichtigung des Cloud Computings: Zur Senkung der Speicher- und Administrationskosten im ILM sollten nicht ausschließlich unternehmensinterne Speichermedien betrachtet werden, da die Verwaltung von Informationen im Cloud Computing ebenfalls zu einer Kostensenkung beitragen kann.⁵⁷³ Das Konzept des Cloud Computings sollte in ein ILM-Vorgehensmodell integriert werden.⁵⁷⁴
- Berücksichtigung rechtlicher Anforderungen: Die SNIA und StorageTek erwähnen in den jeweiligen Vorgehensmodellen lediglich, dass die Informationen mit Zugriffsberechtigungen zu versehen sind, damit ein Zugriff nur durch autorisierte Personen möglich ist. Neben den Zugriffsberechtigungen müssen rechtliche Anforderungen analysiert und berücksichtigt werden, um bspw. Vorschriften zu Aufbewahrungsorten und -fristen bezüglich der Verwaltung von Informationen einzuhalten.⁵⁷⁵

Aufbauend auf den bereits vorgestellten Vorgehensmodellen und der Literaturanalyse wird ein ILM-Vorgehensmodell konstruiert, das aus vier Phasen⁵⁷⁶ besteht.⁵⁷⁷ Die einzelnen Phasen können mehrfach durchlaufen werden.⁵⁷⁸ Die „Analyse der IS-Architektur“, „Beschreibung relevanter Datenbestände“, „Klassifizierung der Daten“

⁵⁷² Ohne dabei vorwiegend den Fokus auf die Verwaltung und den Vertrieb von Speichermedien und Speicherumgebungen zu legen.

⁵⁷³ Vgl. Kapitel 1.3; Kapitel 3

⁵⁷⁴ Vgl. Kapitel 2.2

⁵⁷⁵ Vgl. Kapitel 2.4.3, Kapitel 3.4.2

⁵⁷⁶ Durch die Unterteilung des Modells in einzelne Phasen kann es direkt mit bereits vorhandenen ILM-Vorgehensmodellen verglichen werden. Vgl. zu den Grundsätzen der Klarheit, Vergleichbarkeit und des systematischen Aufbaus bei der ordnungsgemäßen Modellierung Becker, Probandt, Vering / Grundsätze ordnungsgemäßer Modellierung / 35-36.

⁵⁷⁷ Abgeleitet aus den beschriebenen Vorgehensmodellen in Kapitel 5.1.1, welche ebenfalls in einzelne Phasen gegliedert sind.

⁵⁷⁸ Vgl. hierzu die Iteration der Phasen in den Vorgehensmodellen BITKOM und EMC.

und „Verlagerung der Daten“ bilden die einzelnen Phasen und stellen zugleich die wesentlichen Teilaufgaben des ILM dar.⁵⁷⁹ Die Teilaufgaben des ILM werden dabei nicht allein aus den einzelnen Phasen des Informationslebenszyklus abgeleitet⁵⁸⁰ und zielen auch nicht vorwiegend auf die Verwaltung von Speichermedien ab⁵⁸¹. Vielmehr beschreiben die Phasen das allgemeine Vorgehen bei der Umsetzung von ILM.⁵⁸² Begonnen wird mit der „Analyse der IS-Architektur“ und der „Beschreibung relevanter Datenbestände“. Die Anwendungssysteme, Datenbestände, Anwender und Administratoren sowie rechtliche Anforderungen werden ermittelt und untersucht. Da die Klassifizierung wesentlicher Bestandteil von ILM ist,⁵⁸³ erfolgt in der dritten Phase die „Klassifizierung der Daten“. Es wird erörtert, wie der Wert der Informationen mit Hilfe des Nutzungsgrads⁵⁸⁴ ermittelt werden kann und wie die Informationen klassifiziert werden können.⁵⁸⁵ Nach der Klassifizierung der Informationen können diese in der Phase „Verlagerung der Daten“ auf die jeweils relevanten Speicherebenen verlagert werden.⁵⁸⁶ Die Nutzung der Dienste des Cloud Computings wird im Vorgehensmodell berücksichtigt. Mit Hilfe des konstruierten Vorgehensmodells wird in dieser Arbeit die Wirtschaftlichkeit der automatisierten Verwaltung unstrukturierter Daten im ILM analysiert. Wenn eine Anpassung des Vorgehensmodells im Sinne der Wiederverwendung notwendig sein sollte, können Anpassungsmaßnahmen bzw. sogenannte kompositorische Maßnahmen am Modell vorgenommen werden.⁵⁸⁷ Denkbar wäre, dass im Zuge der Einfüh-

⁵⁷⁹ Es werden nur die Sachverhalte im Vorgehensmodell berücksichtigt, die, basierend auf den Erkenntnissen der Untersuchung der bereits vorhandenen Vorgehensmodelle und der Literaturanalyse, zur Beschreibung der Teilaufgaben des ILM relevant sind. Vgl. hierzu zum Grundsatz der Relevanz bei der ordnungsgemäßen Modellierung Becker, Probandt, Vering / Grundsätze ordnungsgemäßer Modellierung / 33-34.

⁵⁸⁰ Die Phasen in den Vorgehensmodellen von StorageTek und des BITKOM basieren im Wesentlichen auf den einzelnen Zyklen des Informationslebenszyklus, von der Entstehung bis zur Archivierung oder Löschung der Informationen.

⁵⁸¹ In den Vorgehensmodellen von EMC und der SNIA wird der Fokus des ILM und der jeweiligen Teilaufgaben bzw. Phasen auf die Verwaltung und den Support von Speichermedien gelegt.

⁵⁸² Auf dieser Basis wird im späteren Verlauf der Arbeit eine Wirtschaftlichkeitsanalyse zum ILM durchgeführt. Vgl. hierzu Kapitel 5.2.

⁵⁸³ Vgl. Abd-El-Malek et al. / self-storage / 55-62; Bhagwan et al. / Management of Data Storage /; Chen / Information valuation / 135-146; Mesnier et al. / File classification / 44-51; Thome, Sollbach / Modelle des ILM / 22-30; Yates-Mercer, Bawden / Managing the paradox / 19-29

⁵⁸⁴ Vgl. Kapitel 2.4.1

⁵⁸⁵ Vgl. Kapitel 5.2.2

⁵⁸⁶ Vgl. hierzu die Einteilung der Informationen und Speichermedien in Klassen bei den vorgestellten Vorgehensmodellen in Kapitel 5.1.1.

⁵⁸⁷ Vgl. Fettke, Loos / Referenzmodelle / 13

rung von ILM in einem Unternehmen die Teilphasen der Phase „Verlagerung der Daten“ bzw. die einzelnen Speicherebenen angepasst werden müssen. Sollte bspw. das Cloud Computing in einem Unternehmen nicht eingesetzt werden,⁵⁸⁸ so kann die Teilphase „Verlagerung zur Cloud-Ebene“ entfernt werden.⁵⁸⁹ Das konstruierte ILM-Vorgehensmodell ist in Abbildung 5-5 dargestellt. Die einzelnen Phasen und Teilphasen des Vorgehensmodells können mehrfach durchlaufen werden.⁵⁹⁰ Ein mehrfaches Durchlaufen der Phasen ist bspw. sinnvoll, um nach bereits erfolgter Klassifizierung und Verlagerung von Daten die IS-Architektur erneut zu untersuchen, falls diese sich im Laufe der Zeit geändert haben sollte. Weiterhin können Teilphasen mehrfach durchlaufen werden, bspw. im Zuge einer periodisch wiederkehrenden Klassifizierung der Daten und der entsprechenden Erhebung sowie Interpretation und Ergänzung von Messwerten.⁵⁹¹ Gleiches gilt für periodisch wiederkehrende Verlagerungsvorgänge der Daten.⁵⁹² Im Folgenden werden die vier Phasen des Vorgehensmodells, „Analyse der IS-Architektur“, „Beschreibung relevanter Datenbestände“, „Klassifizierung der Daten“ und „Verlagerung der Daten“ sowie die jeweiligen Teilphasen näher vorgestellt.

⁵⁸⁸ Bspw. aufgrund rechtlicher Anforderungen. Vgl. hierzu Szenario 1, Kapitel 5.2.3.1.1.1.1.

⁵⁸⁹ Vgl. hierzu eine Fallstudie zur Verwendung des ILM-Vorgehensmodells zur Verwaltung strukturierter Daten in einem Unternehmen der Energiebranche Heinrich, Stelzer / Informationsmanagement / 525-536; Oßmann / Daten ILM / 23-24; Kapitel 4.3.

⁵⁹⁰ Dargestellt durch die Pfeilverbindungen der vier Phasen in Abbildung 5-1. Es wird unterstellt, dass durch die Verbindung der Phasen gleichzeitig eine Iteration der einzelnen Teilphasen möglich ist. Die Teilphasen sind direkter Bestandteil der jeweiligen Phasen und können innerhalb einer Phase mehrfach durchlaufen werden. Dies gilt auch für einzelne Teilphasen.

⁵⁹¹ Vgl. Kapitel 5.1.2.3

⁵⁹² Vgl. Kapitel 5.1.2.4

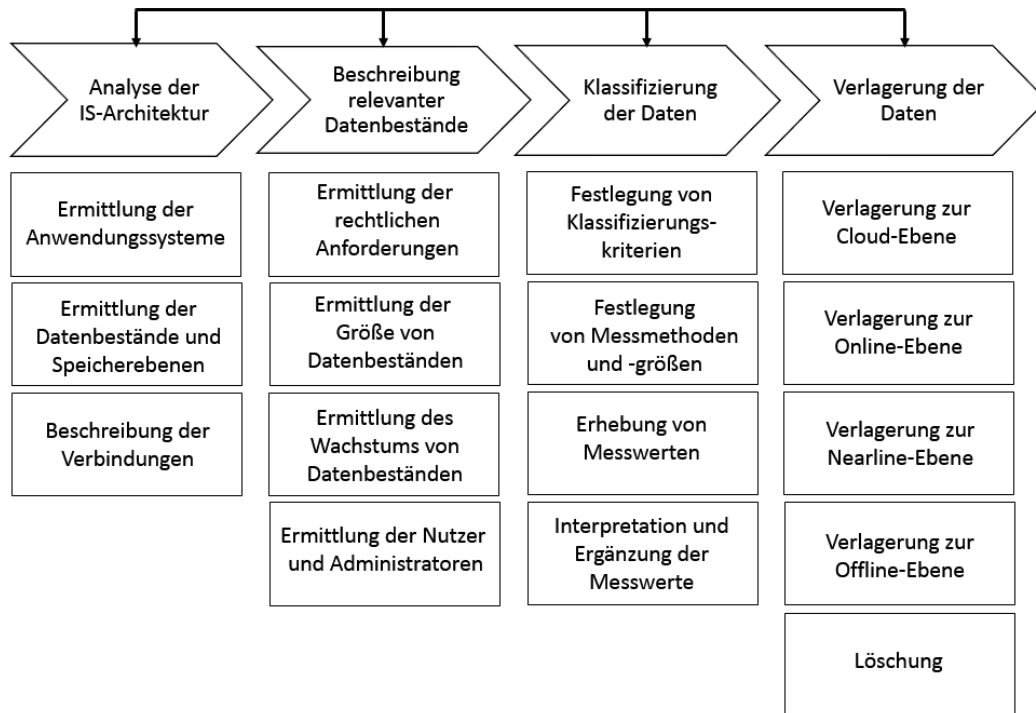


Abb. 5-5: Vorgehensmodell des ILM⁵⁹³

5.1.2.1 Analyse der IS-Architektur

In Phase eins, „Analyse der IS-Architektur“⁵⁹⁴, werden in der Teilphase „Ermittlung der Anwendungssysteme“ zunächst die in einem Unternehmen vorhandenen Anwendungssysteme ermittelt und untersucht. Dabei sind insbesondere die durch die Systeme verarbeiteten Informationen sowie die Verbindungen⁵⁹⁵ zwischen den Systemen von Interesse, die in den Teilphasen „Ermittlung der Datenbestände und Speicherebenen“ und „Beschreibung der Verbindungen“ identifiziert werden. Als typische Beispiele für die zur IS-Architektur gehörigen Anwendungssysteme sind operative Datenbanken, Data-Warehouse-Systeme oder Content-Management-Systeme zu nennen. Durch die Ermittlung von Datenbeständen können Aussagen dazu getroffen werden, welche Informationen in den verschiedenen Anwendungssystemen vorliegen.⁵⁹⁶ Operative Datenbanken

⁵⁹³ Vgl. nach Matthesius, Stelzer / Analyse und Vergleich Konzepte im ILM / 472

⁵⁹⁴ Vgl. zur Beschreibung von Informationssystem-Architekturen Mertens et al. / Grundzüge Wirtschaftsinformatik / 184-185.

⁵⁹⁵ Unter „Verbindungen“ sind die Schnittstellen zu verstehen, die für den Austausch der Informationen zwischen den Anwendungssystemen sorgen. Vgl. hierzu Sneed, Baumgartner, Seidl / Systemtest / 169; Starke / Software-Architekturen / 100-101; Wallmüller / Software-Qualitätsmanagement / 169.

⁵⁹⁶ Vgl. Kapitel 2.4.2

und Data-Warehouse-Systeme verarbeiten zum Beispiel strukturierte Daten und Dokumentenmanagementsysteme verarbeiten unstrukturierte Daten⁵⁹⁷ bzw. Dokumente. Durch die Ermittlung der Datenbestände können im Hinblick auf spätere Phasen des Vorgehensmodells beispielsweise Aussagen dazu getroffen werden, wie die Klassifizierung der Daten vollzogen werden kann und welche Mess- bzw. Klassifizierungsmethoden eingesetzt werden können.⁵⁹⁸ Die Ermittlung der Speicherebenen liefert Erkenntnisse darüber, wo die Daten gespeichert sind. Typische Speicherebenen sind unternehmensinterne Speichermedien, Nearline-Storage-Systeme, Archive oder Cloud-Computing-Systeme.⁵⁹⁹ Die Verbindungen zwischen den Systemen liefern Aussagen dazu, welche Systeme Informationen aus anderen Systemen aufnehmen, verarbeiten und erneut an andere Systeme weitergeben. Ein Beispiel hierfür ist die Extraktion operativer Quellsystemdaten und die anschließende Ablage der Daten in der integrierten Datenbank eines Data-Warehouse-Systems, um mit Hilfe der Werkzeuge⁶⁰⁰ des Data Warehouse die Bereitstellung umfassender und konsistenter Informationen für die Entscheidungsunterstützung in Unternehmen zu ermöglichen.⁶⁰¹ Dabei werden häufig zunächst die operativen Quelldaten in das Data Warehouse geladen und im Sinne der Informationsaufbereitung transformiert und mit anderen Daten kombiniert.⁶⁰² Ein weiteres Beispiel ist die Nutzung eines Dokumentenmanagementsystems, das Dokumente sowohl in der operativen Datenbank als auch in angebundenen Nearline-Storage-Systemen und Archiven speichert. Durch die Analyse der Verbindungen wird deutlich, ob bestimmte Informationen bspw. auf unterschiedlichen Systemen gehalten werden und ob eine mehrfache Aufbewahrung dieser Informationen in den folgenden Phasen des Vorgehensmodells notwendig und sinnvoll ist.

⁵⁹⁷ Vgl. Kapitel 2.3; Kapitel 3.3

⁵⁹⁸ Vgl. hierzu zur automatisierten Klassifizierung von strukturierten Daten Kosler, Matthesius, Stelzer / Klassifizierung von Informationen / 129-145

⁵⁹⁹ Vgl. Kapitel 2.4.2, Kapitel 3.4.1, Kapitel 5.1.2.4

⁶⁰⁰ Vgl. Connolly, Begg, Strachan / Database Systems / 922; Gadatsch, Frick / SAP-BI / 164

⁶⁰¹ Vgl. Hannig / Business Intelligence / 9; Inmon / Data Warehouse / 31

⁶⁰² Vgl. Königer, Reithmayer / Management Informationen / 271; Schwarze / Wirtschaftsinformatik / 370; Stahlknecht, Hasenkamp / Wirtschaftsinformatik / 415

5.1.2.2 Beschreibung relevanter Datenbestände

Nachdem die IS-Architektur untersucht wurde, erfolgt in Phase 2 des Vorgehensmodells, der „Beschreibung relevanter Datenbestände“, die „Ermittlung der rechtlichen Anforderungen“, die „Ermittlung der Größe von Datenbeständen“, die „Ermittlung des Wachstums von Datenbeständen“ und die „Ermittlung der Nutzer und Administratoren“. Die Ermittlung der rechtlichen Anforderungen dient insbesondere der Analyse von gesetzlich vorgeschriebenen Aufbewahrungsorten und -fristen.⁶⁰³ Dabei ist ebenfalls zu überprüfen, ob vorhandene Sicherheitskonzepte einen ausreichenden Schutz der Daten gewährleisten.⁶⁰⁴ Die physische Größe⁶⁰⁵ der Datenbestände, die sich in den relevanten Anwendungssystemen befinden, ist ein wichtiger Indikator für die Beantwortung der Frage, welches Anwendungssystem der IS-Architektur, im Hinblick auf eine Kostenreduktion, zuerst im Rahmen von ILM analysiert werden sollte.⁶⁰⁶ Auch die Berücksichtigung der Größe einzelner Dateien innerhalb der Datenbestände ist im Rahmen des ILM-Vorgehensmodells von Relevanz, da bei der Migration bzw. Verlagerung von Dateien auf langsamere Speichermedien mit großen Dateien begonnen werden sollte, um kostenintensivere Speichermedien möglichst schnell zu entlasten.⁶⁰⁷ Weiterhin muss das Wachstum der jeweiligen Datenbestände im Zeitverlauf untersucht werden. Das Wachstum zeigt, welche Datenbestände in welchen Zeitabständen in die Anwendungssysteme geladen werden und erlaubt zudem eine Prognose⁶⁰⁸ über das zukünftige Wachstum sowie die zukünftige physische Größe der Datenbestände. Somit können bereits vor Beginn der Umsetzung eines ILM-Konzepts Aussagen über die zukünftige Kostenentwicklung von Anwendungssystemen getroffen werden. Je größer das künftig zu erwartende Datenvolumen ist, desto mehr Datenhaltungs- und Datensicherungskosten entstehen. Beispiele für schnell wachsende Datenbestände hinsichtlich unstrukturierter Daten sind Verbindungsdaten von Telekommunikationsanbietern, Abrechnungsdaten von Energieversorgern, Video- und Audiodateien von Anbietern sozialer Netzwerke oder

⁶⁰³ Vgl. Kapitel 2.4.3, Kapitel 3.4.2

⁶⁰⁴ Vgl. Kapitel 2.4.3.2 und Kapitel 3.4.2.2. Hierbei sollten zugleich die Zugriffsrechte der Nutzer untersucht werden. Es muss gewährleistet werden, dass die definierten Zugriffsrechte auch nach einer etwaigen Verlagerung von Informationen auf verschiedene Speicherebenen erhalten bleiben.

⁶⁰⁵ Größe der Datenbestände in Gigabyte oder Terabyte.

⁶⁰⁶ Vgl. Schmitz / Migrationsstrategie / 4

⁶⁰⁷ Vgl. Schmitz / Migrationsstrategie / 4

⁶⁰⁸ Vgl. Backhaus et al. / Analysemethoden / 45-116; Schröder / Zeitreihenprognose / 11-14

digitale Bildarchive von Krankenhäusern. Zusätzlich werden die Nutzer und Administratoren der Datenbestände ermittelt, da diese für die dritte Phase des Vorgehensmodells, der „Klassifizierung der Daten“, notwendig sind.

5.1.2.3 Klassifizierung der Daten

In der dritten Phase des ILM-Vorgehensmodells erfolgt die „Klassifizierung der Daten“, um die Datenbestände entsprechend der ermittelten Klasse in der jeweils geeigneten Speicherebene aufzubewahren oder zu löschen.⁶⁰⁹ In der Teilphase „Festlegung von Klassifizierungskriterien“ werden für die Informationsklassifizierung relevante Kriterien festgelegt, wie beispielsweise die Zugriffshäufigkeit auf die Informationen sowie das Anwender- und Administratorenwissen über die Systeme und Informationen.⁶¹⁰ Vorgehensweise und Methoden zur Messung der relevanten Kriterien werden in der Teilphase „Festlegung von Messmethoden und -größen“ bestimmt. Anschließend erfolgt die „Erhebung von Messwerten“ und die „Interpretation und Ergänzung der Messwerte“. In diesen beiden Teilphasen werden die Informationen entsprechend der festgelegten Kriterien und Methoden klassifiziert. Die Klassifizierung kann nicht vollständig automatisiert erfolgen, da neben den technischen Klassifizierungskriterien, wie Zugriffshäufigkeiten und Zugriffszeitpunkten, zusätzliche Informationen, wie Serviceebenen-Vereinbarungen, Aspekte der Wirtschaftlichkeit, rechtliche Anforderungen oder Kenntnisse über die zukünftige Nutzung von Datenbeständen, für die Verlagerung von Daten notwendig sind. Aus diesem Grund müssen die automatisch erhobenen Messwerte interpretiert und ergänzt werden. Eine Klasse bezeichnet die geeignete Speicherebene,⁶¹¹ auf welche die Informationen später verlagert werden sollen. Hierbei werden die Informationen in fünf Klassen eingeteilt:⁶¹²

⁶⁰⁹ Vgl. hierzu Kapitel 2.4.2

⁶¹⁰ Vgl. hierzu Kapitel 2.4.1. Empfehlenswert ist bspw. auch die Berücksichtigung von zeitkritischen Zugriffen auf die Informationen sowie die Notwendigkeit zur Einhaltung rechtlicher Vorschriften. Informationen, die jederzeit einen zeitnahen Zugriff erfordern und zugleich vertrauliche Informationen darstellen, sollten auf unternehmensinternen Speichermedien mit kurzen Antwortzeiten abgelegt werden. Informationen, an die weniger Ansprüche hinsichtlich zeitnaher Zugriffe und rechtlicher Anforderungen gestellt werden, können bspw. auf eine Public Cloud oder ein Archivsystem verlagert werden. Aus Kostengesichtspunkten sollte auch die Größe der Datenmenge berücksichtigt werden. Die Nutzung von Public-Cloud-Diensten ist bspw. aufgrund hoher Zeit- und Kostenaufwände nicht lohnend, wenn oft große Datenmengen in die Public Cloud übertragen werden müssen. Vgl. Kapitel 3.4.1.

⁶¹¹ Vgl. Kapitel 5.1.2.4

⁶¹² Die Einteilung der Klassen, Speicherebenen und die jeweiligen Hinweise zu den typischen Eigenschaften der zu verwaltenden Informationen stellen lediglich Empfehlungen des Verfassers dar. Entsprechend der Unternehmensanforderungen können diese Klassen erweitert und modifiziert werden.

- Klasse „Cloud“: Informationen, auf die oft zugegriffen wird und die aus rechtlicher Sicht ohne Einschränkungen verwaltet werden können, sollten mit Hilfe von Public-Cloud-Diensten verwaltet⁶¹³ werden.⁶¹⁴ Somit können die Administrationskosten für Datenträger gesenkt werden.⁶¹⁵
- Klasse „Online“: Informationen dieser Klasse werden oft von Anwendern benötigt und sind aufgrund der Unternehmensanforderungen und rechtlichen Anforderungen als vertraulich und schutzwürdig einzustufen. Sie sollten von daher für die schnelle und sichere Verwendung und Ablage im Online-Speicherbereich vorgehalten werden.
- Klasse „Nearline“: Vertrauliche Informationen, auf die in regelmäßigen Abständen zugegriffen wird, für die aber eine Ablage auf kostenintensiven Speichermedien nicht lohnend ist, werden auf unternehmensinternen Nearline-Storage-Systemen⁶¹⁶ abgelegt.

⁶¹³ Ebenfalls denkbar ist die Nutzung von Cloud-Computing-Diensten zur Bedienung von Lastspitzen (Vgl. Kapitel 3.1) und zur flexiblen Anpassung der Ressourcen. Vgl. Kapitel 3.2.

⁶¹⁴ Hierfür bietet sich die Nutzung von Speicherkapazitäten oder Anwendungssystemen über Cloud-Dienste an (IaaS und SaaS, Vgl. hierzu Kapitel 3.2). Es existieren auch Cloud-Dienste, in denen nach Anbietervorgaben vertrauliche und unternehmenskritische Daten verwaltet werden können (Vgl. Grobman / ERP-Systeme On Demand / 15). Unternehmen sollten vor der Nutzung derartiger Systeme immer überprüfen, ob der Anbieter Zertifizierungen zur Einhaltung von Sicherheitsstandards vorweisen kann (Vgl. Kapitel 3.4.2.2). Zusätzlich ist von Bedeutung, ob der Anbieter seinen Firmensitz in Deutschland hat und die Daten möglichst innerhalb von Deutschland verwaltet (Vgl. Bundesministerium der Justiz / BDSG / § 1; In diesem Fall gilt das Bundesdatenschutzgesetz). Die Nutzung von Public-Cloud-Diensten erfordert großes Vertrauen in den Anbieter, da der Nutzer kaum Einblick in die Geschäftsprozesse des Anbieters hat (Vgl. Winkler / Securing the Cloud / 55-66). Es kann nicht vollständig davon ausgegangen werden, dass die verwalteten vertraulichen und unternehmenskritischen Informationen frei von Beeinträchtigungen und ohne Risiken für den Nutzer in der Public Cloud verwaltet werden können (Vgl. Levy / NSA killed the Internet / ; Venhaus, Haselbeck, Wintermann / Schatten der Überwachung / 40-50; Wicker / Death of Privacy / 129-134). Aus diesen Gründen ist es empfehlenswert, derartige Daten auf den unternehmenseigenen Speichermedien zu verwalten. Die verwalteten Informationen auf unternehmensinternen Speichermedien sind zwar ebenfalls Beeinträchtigungen, wie Datenverlust, Verfälschung der Daten oder Spionage, ausgesetzt, allerdings hat der Nutzer hier einen deutlich größeren Einfluss auf den Schutz der Daten.

⁶¹⁵ Vgl. Schadler / Email in the Cloud / 3-5

⁶¹⁶ Vgl. Liu, Li, Zhang / Dispatching Methods / 690-695; Maier, Hädrich, Peinl / Knowledge Infrastructures / 254-257

- Klasse „Offline“: Informationen, auf die nur in Ausnahmefällen zugegriffen wird oder deren langfristige Aufbewahrung aufgrund rechtlicher Anforderungen notwendig ist, können auf kostengünstigen Speichermedien aufbewahrt bzw. archiviert⁶¹⁷ werden.⁶¹⁸
- Klasse „Löschen“: Auf Informationen dieser Gruppe wird nicht mehr zugegriffen. Sie haben keinen Wert für das Unternehmen und können gelöscht werden, sofern keine rechtlichen Aufbewahrungsrestriktionen bestehen.

5.1.2.4 Verlagerung der Daten

In Phase vier, der „Verglagerung der Daten“, werden die Informationen auf die jeweils dafür vorgesehene Speicherebene verglagert. Die Notwendigkeit zur Verglagerung der Daten ergibt sich aus den Ergebnissen der Klassifizierung der Daten. Eine Beschreibung des Vorgehens zur Klassifizierung und anschließenden Verglagerung der Daten erfolgt in Kapitel 5.2.2. Zur Senkung der Administrationskosten werden die Daten mit Hilfe des Cloud Computings verwaltet („Verglagerung zur Cloud-Ebene), wenn dem keine Anforderungen aus Sicht der IT-Sicherheit oder aus rechtlicher Sicht⁶¹⁹ entgegenstehen. Haben die Informationen bspw. einen sehr hohen Nutzungsgrad und unterliegen rechtlichen Anforderungen, so verbleiben die Daten in dem jeweiligen System oder werden zurück in das System verglagert („Verglagerung zur Online-Ebene“). Informationen, auf die in regelmäßigen Abständen zugegriffen werden muss, für die aber eine Speicherung im operativen System wirtschaftlich nicht sinnvoll ist, werden auf so genannten Nearline-Storage-Systemen⁶²⁰ (NLS) abgelegt („Verglagerung zur NLS-Ebene“). Die Informationen werden auf Archivierungssysteme verglagert, wenn auf diese für einen längeren Zeitraum nicht zugegriffen werden muss⁶²¹ und die Aufbewahrung aber, zum Beispiel aufgrund rechtlicher Anforderungen, notwendig ist („Verglagerung zur Offline-Ebene“). Nicht mehr benötigte Informationen werden gelöscht („Löschen“).

⁶¹⁷ Vgl. Küspert, Schaarschmidt, Herbst / Archiv Transaktionen / 195-211

⁶¹⁸ Hierfür können Magnetbänder verwendet werden. Vergleich hierzu Janko / Informationswirtschaft / 112. Die Speicherung auf Magnetbändern gilt als kostengünstigste Alternative. Vergleich hierzu Hillyer, Silberschatz / Tertiary storage systems / 195-204.

⁶¹⁹ Vgl. Kapitel 3.4.2

⁶²⁰ Vgl. Liu, Li, Zhang / Dispatching Methods / 690-695; Maier, Hädrich, Peinl / Knowledge Infrastructures / 109

⁶²¹ Vgl. Küspert, Schaarschmidt, Herbst / Archiv Transaktionen / 195-211

5.1.3 Automatisierte Informationsklassifizierung im Rahmen des ILM

In Kapitel 5.1.2 wurde ein Vorgehensmodell zum ILM konstruiert. In der dritten Phase des Vorgehensmodells wurde die Teilphase „Klassifizierung der Daten“ beschrieben, in welcher die Datenbestände entsprechend ihres Nutzungsgrads klassifiziert werden.⁶²² Als schwierig erweist sich die Wertzuweisung und anschließende Klassifizierung von Informationen.⁶²³ Hierfür werden zahlreiche Angaben bzw. Klassifizierungskriterien benötigt, wie beispielsweise die Anzahl der Lese- und Schreibzugriffe auf diese Objekte.⁶²⁴ Die Erhebung, kontinuierliche Aktualisierung und Auswertung dieser Klassifizierungskriterien ohne automatisierte Unterstützung ist praktisch kaum möglich und wirtschaftlich nicht sinnvoll. Aus diesem Grund muss die Informationsklassifizierung zumindest teilweise automatisiert erfolgen. In der Literatur finden sich verschiedene Konzepte zur Gestaltung der automatisierten Informationsklassifizierung im ILM.⁶²⁵ Diese Konzepte wurden durch die Analyse von Zeitschriften und Konferenzen sowie die Suche mit Suchmaschinen-Diensten identifiziert.⁶²⁶ Die Konzepte werden im Folgenden vorgestellt und miteinander verglichen. Auf Basis der Untersuchung dieser Konzepte werden Anforderungen an die automatisierte Informationsklassifizierung definiert.⁶²⁷

5.1.3.1 Vorstellung und Vergleich von Konzepten zur automatisierten Klassifizierung unstrukturierter Daten

5.1.3.1.1 Vorstellung der Konzepte

Chen verwendet zur automatisierten Klassifizierung unstrukturierter Daten den Nutzungsgrad als Klassifizierungskriterium.⁶²⁸ Als wesentliche Anforderungen nennt Chen die Verwendung messbarer Indikatoren zur Klassifizierung der Informationen sowie die Verringerung der Notwendigkeit von Anwenderinteraktionen. Basierend auf der Zu-

⁶²² Vgl. hierzu Kapitel 5.1.2.3

⁶²³ Vgl. Abd-El-Malek et al. / self-storage / 55-62; Bhagwan et al. / Management of Data Storage /; Chen / Information valuation / 135-146; Mesnier et al. / File classification / 44-51; Thome, Sollbach / Modelle des ILM / 22-30; Yates-Mercer, Bawden / Managing the paradox / 19-29

⁶²⁴ Vgl. Chen / Information valuation / 135-146; Ellard et al. / File prediction / 1-14

⁶²⁵ Vgl. Chen / Information valuation / 135-146; Bhagwan et al. / Management of Data Storage /; Chandra, Gehani, Yu / Automated Storage / 12; Mesnier et al. / File classification / 44-51; Shah et al. / Classification for ILM / 1-15; Verma et al. / Very Large File Systems / 160-168; Zadok et al. / Reducing Storage Management Costs / 101-105

⁶²⁶ Vgl. Kapitel 1.3.2

⁶²⁷ Vgl. hierzu Kapitel 5.1.3.2

⁶²⁸ Vgl. Chen / Information valuation / 135-146; Vgl. Kapitel 2.4.1

griffshistorie einer Information wird ihre Klassenzugehörigkeit für einen Zeitpunkt t berechnet. Die Zugriffshäufigkeiten der Vergangenheit werden dabei in verschiedene Zyklen aufgeteilt. Anhand dieser Zyklen wird mit Hilfe der Klassifizierungsfunktion die Klasse einer Information berechnet, wobei aktuelle Zyklen stärker gewichtet werden als ältere. Je häufiger in der Vergangenheit auf eine Information zugegriffen wurde und je öfter diese Zugriffe kurz vor dem Klassifizierungszeitraum t erfolgten, desto höher der Wert, der den Informationen durch die Klassifizierung zugewiesen wird. Die Informationen werden entsprechend ihres Nutzungsgrads klassifiziert.⁶²⁹ Ein Vorteil des Konzepts von Chen ist die Verwendung des Nutzungsgrads als Klassifizierungskriterium, da sich dieses ohne großen Aufwand ermitteln und messen lässt. Auch die höhere Gewichtung aktuellerer Zugriffszyklen als ältere erweist sich als sinnvoll.⁶³⁰ Nachteilig ist die Vernachlässigung des Anwender- und Administratorenwissens über die Systeme und Informationen. Die Informationsklassifizierung basiert ausschließlich auf technischen Kriterien. Die prototypische Umsetzung des Konzepts erfolgt auf Basis von Datenbeständen mit einem physischen Volumen von 48 Gigabyte. Das Ziel des Konzepts ist die Reduktion von Speicher- und Administrationskosten. Das Kosten- und Nutzenverhältnis und inwiefern die Implementierung und die Nutzung wirtschaftlich sinnvoll sind, wird innerhalb des Konzepts nicht untersucht.

Shah et al. und Alatorre et al. klassifizieren Informationen ebenfalls anhand des Nutzungsgrads.⁶³¹ Im Gegensatz zum Konzept von Chen basiert die Klassifizierung bei diesen beiden Konzepten nicht ausschließlich auf technischen Kriterien. Anwender und Administratoren wirken aktiv an der Gestaltung der Klassifizierungsfunktionen und der anschließenden Verlagerung der Informationen mit. Neben der Berücksichtigung des Anwender- und Administratorenwissens stehen in den Konzepten von Shah et al. und Alatorre et al. die Systemleistung und die Reduktion der Kosten für Datenspeicher im Vordergrund.⁶³² Shah et al. untersuchen einen Datenbestand mit einer physischen Größe

⁶²⁹ Die Klassifizierung erfolgt dabei auf einer Werteskala von 1 (unwichtig) bis 10 (sehr wichtig).

⁶³⁰ Somit wird beispielsweise vermieden, dass für Informationen, die zwar für ältere Zyklen hohe Zugriffe aufweisen, jedoch aktuell kaum eine Rolle spielen, überbewertet werden. Diese Vorgehensweise ist sinnvoll, wenn die Klassifizierung der Informationen in größeren Zeitabständen, bspw. jährlich, stattfindet.

⁶³¹ Vgl. Shah et al. / Classification for ILM / 1-15; Alatorre et al. / Intelligent ILM / 9-18

⁶³² Ziel ist die Reduktion von Kosten. Die Verfasser treffen jedoch keine Aussagen zur Wirtschaftlichkeit des Konzepts.

von 1,4 Terabyte. Bei der Klassifizierung wird im Konzept von Shah et al. nicht jede einzelne Information durch die Klassifizierungsfunktion untersucht und klassifiziert, so wie es beim Konzept von Chen der Fall ist. Es werden lediglich prozentuale Anteile von unstrukturierten Daten in Dateiordnern klassifiziert oder Stichproben bzw. kleine Datenbestände aus dem gesamten Datenbestand gezogen. Die Klassifizierung wird damit vereinfacht, indem die Klassifizierungsergebnisse der ausgewählten Daten auf die übrigen Daten übertragen werden. Die Klassifizierung wird somit ungenauer, allerdings wird die Leistungsfähigkeit der genutzten Hardware aufrechterhalten.⁶³³ Alatorre et al. führen die Klassifizierungstests mit Datenmengen von bis zu 8 Terabyte durch. Bei der Klassifizierung können einzelne Dateien oder gesamte Speichermedien mit Hilfe der Klassifizierungsfunktionen klassifiziert werden.

Bhagwan et al. verfolgen ein ähnliches Konzept zur Klassifizierung von Informationen.⁶³⁴ Ähnlich wie beim Konzept von Chen spielt das Anwender- und Administratorenwissen eine untergeordnete Rolle - die Klassifizierung anhand technischer Kriterien steht im Vordergrund. Durch Verwendung des Kriteriums Nutzungsgrad werden Informationen bewertet und klassifiziert. Im Vordergrund stehen die Leistung der Systeme sowie die Kostenreduktion, die durch eine optimale Verlagerung der unstrukturierten Daten erreicht werden kann. Im Konzept von Bhagwan et al. wird nicht jede einzelne Information untersucht und klassifiziert. Es wird empfohlen, zusammengehörige Daten vor der automatisierten Klassifizierung in Dateiordnern zu speichern. Die Klassifizierungsfunktionen klassifizieren den Nutzungsgrad des entsprechenden Dateiordners. Auf diese Weise wird die Klassifizierung der Daten vereinfacht. Eine prototypische Implementierung des Konzepts erfolgte nicht.

Verma et al.⁶³⁵ und Mesnier et al.⁶³⁶ verwenden in ihren Konzepten zur automatisierten Klassifizierung von unstrukturierten Daten anstelle des Klassifizierungskriteriums Nutzungsgrad den Typ der Information bzw. den Dateityp. Informationen gleichen Typs, wie beispielsweise *.txt, besitzen nach Meinung der Autoren dieselben Eigenschaften und können einer gemeinsamen Klasse zugeordnet werden. Dabei wird überprüft, ob

⁶³³ Vgl. hierzu Shah et al. / Classification for ILM / 12-14

⁶³⁴ Vgl. Bhagwan et al. / Management of Data Storage /

⁶³⁵ Vgl. Verma et al. / Very Large File Systems / 160-168

⁶³⁶ Vgl. Mesnier et al. / File classification / 44-51

zum Beispiel eine *.txt-Datei größer als ein Megabyte ist und ob die Auslastung des Speichermediums größer als achtzig Prozent ist. Trifft beides zu, wird die *.txt-Datei auf ein kostengünstigeres Speichermedium verlagert. Mesnier et al. unterscheiden zusätzlich, ob auf eine Information ein schreibender oder ein lesender Zugriff erfolgt. Auf Informationen vom Typ *.log erfolgt in der Regel ein schreibender Zugriff, weshalb diese Information entsprechend klassifiziert und danach auf Speichermedien verlagert wird, die für den Schreibzugriff optimiert sind. Allerdings vernachlässigen die Konzepte von Verma et al. und Mesnier et al. den Nutzungsgrad der Informationen. Somit könnten Informationen auf kostenintensive Speichermedien verlagert werden, obwohl seit einem längeren Zeitraum keine Zugriffe darauf erfolgten und eine Speicherung auf kostengünstigen Speichermedien sinnvoller wäre. Die prototypische Umsetzung der Konzepte von Verma et al. und Mesnier et al. erfolgte auf Basis von Datenbeständen mit einem physischen Volumen von 40 Gigabyte bzw. 53 Gigabyte. Die Betrachtung der Wirtschaftlichkeit erfolgte nicht.

Eine starke Konzentration auf das Anwender- und Administratorenwissen bezüglich der Systeme und Informationen erfolgt in den Konzepten von Zadok et al.⁶³⁷ sowie Chandra, Gehani und Yu⁶³⁸. Jeder Anwender und Administrator erzeugt eigene Klassifizierungsfunktionen bzw. klassifiziert seine Informationen selbst. Dabei wird festgelegt, für welchen Zeitraum eine Information einen hohen Klassifizierungswert besitzt und ab wann eine Information auf kostengünstigere Speichermedien verlagert werden kann oder gelöscht werden soll. Dies führt zu Kostensenkungen, da Informationen entsprechend ihres zugewiesenen Werts verlagert werden. Allerdings setzen die drei Konzepte eine angemessene Vorgehensweise der Anwender und Administratoren voraus. Denkbar ist aber, dass aus subjektiver Sicht der Anwender und Administratoren alle Informationen wichtig sind, obwohl nur selten darauf zugegriffen wird.⁶³⁹ Die Verwendung technischer und objektiver Klassifizierungskriterien wird bei diesen Konzepten vernachlässigt. Dies führt auch dazu, dass für die Informationsklassifizierung ein hoher Aufwand entsteht, da mögliche Kostensenkungspotentiale, durch eine automatisierte Unterstüt-

⁶³⁷ Vgl. Zadok et al. / Reducing Storage Management Costs / 101-105

⁶³⁸ Vgl. Chandra, Gehani, Yu / Automated Storage / 12

⁶³⁹ Vgl. Chen / Information valuation / 135-146

zung der Klassifizierung, nicht genutzt werden.⁶⁴⁰ Zadok et al. setzen die prototypische Umsetzung ihres Konzepts auf einem Datenbestand mit einem physischen Volumen von 30 Gigabyte um und Chandra, Gehani und Yu auf Datenvolumina von 80 und 120 Gigabyte.

Turczyk legt in seinem Konzept den Schwerpunkt der Untersuchungen auf die Prognose der Wahrscheinlichkeit zukünftiger Zugriffe im Zeitverlauf. Je höher die Wahrscheinlichkeit künftiger Zugriffe im Zeitverlauf ist, desto höher ist der Wert, der den Informationen zugewiesen wird. Basierend auf der Historie des Nutzungsgrads werden die zukünftigen Zugriffshäufigkeiten und Zugriffszeitpunkte prognostiziert bzw. berechnet. Damit wird die automatisierte Informationsklassifizierung vereinfacht, da Zugriffshäufigkeiten und Zugriffszeitpunkte sowie Klassenzugehörigkeiten nicht ständig neu berechnet werden müssen, sondern eine Vorausplanung bezüglich eines definierten Zeitraums erfolgen kann.⁶⁴¹ Dies wirkt sich positiv auf die Leistungsfähigkeit der in den Unternehmen vorhandenen Systeme und damit auf die Kostensenkungspotentiale aus. Allerdings werden die Klassifizierungsergebnisse ungenauer, da durch die Prognose lediglich Schätzungen bezüglich des zukünftigen Nutzungsgrads angestellt werden können.⁶⁴² Das Anwender- und Administratorenwissen wird berücksichtigt, da eine anwenderspezifische Anpassung der Methoden der Wertzuweisung innerhalb des Konzepts möglich ist. Die prototypische Umsetzung des Konzepts erfolgt auf einer physischen Datenbasis mit einem Volumen von 2 Gigabyte und 1.000 Dateien. Das Kosten- und Nutzenverhältnis bzw. die Wirtschaftlichkeit hinsichtlich Implementierung und Einsatz des Konzepts werden nicht untersucht.⁶⁴³

5.1.3.1.2 Vergleich der vorgestellten Konzepte

In Abbildung 5-6 werden die vorgestellten Konzepte gegenübergestellt. Es wird gezeigt, welche Anforderungen die Verfasser an ihre jeweiligen Konzepte stellen. Auf dieser Basis und unter Berücksichtigung der in Kapitel 2.4 definierten Anforderungen an ILM

⁶⁴⁰ Der Aspekt der Wirtschaftlichkeit wird in den Konzepten von Zadok et al. sowie Chandra, Gehani und Yu nicht betrachtet.

⁶⁴¹ Vgl. Abd-El-Malek et al. / self-storage / 55-62; Douglass et al. / Delete-Optimized Storage /; Mesnier et al. / File classification / 44-51; Shah et al. / Classification for ILM / 1-15

⁶⁴² Vgl. Bhagwan et al. / Management of Data Storage /; Shah et al. / Classification for ILM / 12-14; Turczyk / Wertzuweisung / 462

⁶⁴³ Vgl. Turczyk / ILM / 117-121

werden in Kapitel 5.1.3.2 Anforderungen an die automatisierte Informationsklassifizierung definiert.

Konzept \ Anforderungen	Chen 2005	Shah et al. 2006	Alatorre et al. 2014	Bhagwan et al. 2005	Verma et al. 2005	Mesnier et al. 2004	Zadok et al. 2004	Chandra, Gehani, Yu 2007	Turezyk 2008
Nutzungsgrad									
Nutzungsgrad als Klassifizierungskriterium	X	X	X	X					X
Klassifizierung									
Einteilung Informationen in Klassen	X	X	X	X	X	X	X	X	X
Rechtliche Anforderungen									
Berücksichtigung rechtlicher Anforderungen									
Automatisierung									
Klassifizierung der Informationen	X	X	X	X	X	X	X	X	X
Berücksichtigung Anwender- und Administratorenwissen		X	X		X	X	X	X	X
Prognose des Nutzungsgrads									X
Kosten und Wirtschaftlichkeit									
Berücksichtigung Leistungsfähigkeit der Systeme		X	X	X	X				X
Ziel der Kostenreduktion	X	X	X	X	X	X	X	X	X
Berücksichtigung Cloud Computing als Speicherebene									
Untersuchung der Wirtschaftlichkeit									
Datenmengen									
Berücksichtigung großer Datenmengen		X	X						

Abb. 5-6: Vergleich der Konzepte⁶⁴⁴

Alle vorgestellten Konzepte verwenden Klassifizierungsfunktionen für die automatisierte Informationsklassifizierung.⁶⁴⁵ Nach der Wertzuweisung werden die Informationen klassifiziert⁶⁴⁶, um diese anschließend auf die jeweils sinnvollste Speicherebene zu verlagern. Eine Reduktion der Kosten für die Administration der Datenspeicher und die Aufbewahrung der Informationen ist ebenfalls Ziel dieser Konzepte.⁶⁴⁷ Die Nutzung des Cloud Computings als Speicherebene⁶⁴⁸ und die Wirtschaftlichkeit der automatisierten Klassifizierung wird in keinem der Konzepte berücksichtigt bzw. untersucht.⁶⁴⁹ Die Mehrzahl der Konzepte lässt das Anwender- und Administratorenwissen in die Klassifi-

⁶⁴⁴ In Anlehnung an Matthesius, Stelzer / Analyse und Vergleich Konzepte im ILM / 478

⁶⁴⁵ Vgl. Abbildung 5-6, Anforderung Automatisierung (Klassifizierung der Informationen)

⁶⁴⁶ Vgl. Abbildung 5-6, Anforderung Klassifizierung (Einteilung Informationen in Klassen)

⁶⁴⁷ Vgl. Abbildung 5-6, Anforderung Kosten und Wirtschaftlichkeit (Ziel der Kostenreduktion)

⁶⁴⁸ Vgl. Kapitel 5.1.2.3

⁶⁴⁹ Vgl. Abbildung 5-6, Anforderung Kosten und Wirtschaftlichkeit (Berücksichtigung Cloud Computing als Speicherebene); Abbildung 5-6, Anforderung Kosten und Wirtschaftlichkeit (Untersuchung der Wirtschaftlichkeit)

zierung der Informationen einfließen.⁶⁵⁰ Das Klassifizierungskriterium Nutzungsgrad sowie die Leistungsfähigkeit der Systeme werden von fünf der insgesamt neun betrachteten Konzepte berücksichtigt.⁶⁵¹ Eines der vorgestellten Konzepte verwendet die Prognose des Nutzungsgrads.⁶⁵² Keines der Konzepte berücksichtigt die rechtlichen Anforderungen bzw. lässt diese in die Informationsklassifizierung einfließen.⁶⁵³ Dieses Gebiet ist ein offenes Forschungsthema.⁶⁵⁴ Ansätze hierfür liefern Park und Mont.⁶⁵⁵ Sie definieren Funktionen, die ein Ziel, ein Ereignis und eine Aktion beinhalten. Das Ziel beschreibt das Speichermedium, auf das die Informationen verlagert werden. Das Ereignis enthält die Vorgehensweise, wie die Informationen aufgrund rechtlicher Anforderungen behandelt werden sollen, beispielsweise ab wann vertrauliche Informationen gelöscht werden müssen. Die Aktion beschreibt, ob die Informationen verlagert, kopiert oder gelöscht werden müssen. Eine automatisierte Aufnahme und Anpassung rechtlicher Anforderungen ist dennoch nicht möglich. Diese Aufgabe obliegt den Anwendern und Administratoren.⁶⁵⁶ Weiterhin ist zu beachten, dass der Aufwand für den Einsatz von ILM bei kleinen Datenbeständen in keinem angemessenen Verhältnis zur Kostensenkung steht.⁶⁵⁷ In zwei der betrachteten Konzepte wurde prototypisch aufgezeigt, dass die automatisierte Klassifizierung bei großen Datenbeständen praktikabel einsetzbar ist.⁶⁵⁸ Die automatisierte Erhebung, Aktualisierung und Auswertung der Klassifizierungskriterien wird in den meisten Konzepten vereinfacht, indem nicht jedes einzelne

⁶⁵⁰ Vgl. Abbildung 5-6, Anforderung Automatisierung (Berücksichtigung Anwender- und Administratorenwissen)

⁶⁵¹ Vgl. Abbildung 5-6, Anforderung Nutzungsgrad (Nutzungsgrad als Klassifizierungskriterium); Abbildung 5-6, Anforderung Kosten und Wirtschaftlichkeit (Berücksichtigung Leistungsfähigkeit der Systeme)

⁶⁵² Vgl. Abbildung 5-6, Anforderung Automatisierung (Prognose des Nutzungsgrads)

⁶⁵³ Vgl. Abbildung 5-6, Rechtliche Anforderungen (Berücksichtigung rechtlicher Anforderungen)

⁶⁵⁴ Vgl. Vgl. Hadjiantonis, Pavlou / Self-Management / 257-258; Krishnan / Big Data / 113-114; Shen et al. / Data Protection in the Cloud / 137-139; Mont et al. / Privacy Policy Enforcement / 138; Abd-El-Malek et al. / self-storage / 55-62; Johnson, Agrawal / Law and Technology / ; Mont, Beato / Obligation Policies / 51-55

⁶⁵⁵ Vgl. Park / Privacy Policy Grid / 91-100; Mont et al. / Privacy Policy Enforcement / 138

⁶⁵⁶ Aus diesem Grund sollten Konzepte zur automatisierten Klassifizierung nur bedingt eingesetzt werden, bspw. um automatisiert die Zugriffshäufigkeiten oder zukünftige Zugriffe auf die Daten zu ermitteln. Vor der Verlagerung der Daten zwischen den Speicherebenen sollte durch Experten und Administratoren überprüft werden, ob rechtliche Anforderungen beachtet werden müssen oder der Transfer großer Datenmengen zwischen den Speicherebenen wirtschaftlich sinnvoll ist (Vgl. Kapitel 2.4.3, Kapitel 3.4.2, Kapitel 5.1.2.3).

⁶⁵⁷ Vgl. Heinrich, Stelzer / Informationsmanagement / 275

⁶⁵⁸ Vgl. Abbildung 5-6, Anforderung Datenmengen (Berücksichtigung großer Datenmengen)

Datenobjekt zur Klassifizierung herangezogen wird. Es werden Anteile von Informationen in Dateiordnern klassifiziert oder Stichproben bzw. kleine Datenbestände aus dem gesamten Datenbestand gezogen.⁶⁵⁹ Durch diese Vorgehensweise wird die Leistungsfähigkeit der Systeme positiv beeinflusst. Allerdings werden die Klassifizierungsergebnisse ungenauer, da nicht jede Information in die automatisierte Klassifikation einbezogen wird und die Besonderheiten jeder einzelnen Information nicht berücksichtigt werden.⁶⁶⁰ Die Folge ist, dass nicht jede Information auf die jeweils sinnvollste Speicherebene verlagert werden kann. Zugleich sind nur grobe Schätzungen über den künftigen Speicherbedarf und der damit einhergehenden Speicher- und Administrationskosten möglich.

5.1.3.2 Anforderungen an Konzepte zur automatisierten Klassifizierung unstrukturierter Daten

5.1.3.2.1 Verwendung des Nutzungsgrads als Klassifizierungskriterium

Zunächst sollten Konzepte zur automatisierten Klassifizierung von Informationen die Kriterien Zugriffshäufigkeit und Zugriffszeitpunkt zur Bestimmung des Nutzungsgrads der Informationen verwenden.⁶⁶¹ Ein Zugriff erfolgt, wenn eine Information gelesen oder verändert wird.⁶⁶² Die Zugriffshäufigkeit variiert im Zeitverlauf.⁶⁶³ Die häufigsten Zugriffe erfolgen erfahrungsgemäß kurz nach dem Entstehen der Information.⁶⁶⁴ Danach wird in der Regel nur noch selten auf die Information zugegriffen. Etwa sechzig bis achtzig Prozent der Informationen in einem Unternehmen bleiben ungenutzt. Auf sie wird nicht mehr zugegriffen.⁶⁶⁵ Weiterhin muss ermittelt werden, zu welchen Zeitpunkten auf eine Information zugegriffen wurde, um zu bestimmen, ob Informationen vor

⁶⁵⁹ Vgl. Shah et al. / Classification for ILM / 9-14; Turczyk / ILM / 90, 101-114, 140; Verma et al. / Very Large File Systems / 160-168

⁶⁶⁰ Vgl. Bhagwan et al. / Management of Data Storage / ; Shah et al. / Classification for ILM / 12-14; Turczyk / Wertzuweisung / 462

⁶⁶¹ Vgl. Kapitel 2.4.1, Kapitel 2.4.2

⁶⁶² Vgl. Shah et al. / Classification for ILM / 1-15

⁶⁶³ Vgl. Bhagwan et al. / Management of Data Storage / ; Chen / Information valuation / 135-146; Shah et al. / Classification for ILM / 1-15

⁶⁶⁴ Vgl. Chen / Information valuation / 135-146; Zadok et al. / Reducing Storage Management Costs / 101-105).

⁶⁶⁵ Vgl. Bennett, Bauer, Kinchlea / NFS environments / 18-25; Chen / Information valuation / 135-146; Inmon / ILM Data Warehousing / 1-20; Shah et al. / Classification for ILM / 1-15; Short / ILM / 2-37; Zadok et al. / Reducing Storage Management Costs / 101-105

kürzerer Zeit genutzt wurden oder ob für einen längeren Zeitraum nicht darauf zugegriffen worden ist.⁶⁶⁶

5.1.3.2.2 Implementierung geeigneter Klassifizierungsfunktionen

Für die Automatisierung müssen Funktionen implementiert werden, welche die Informationen anhand der Klassifizierungskriterien bewerten und anschließend den zugehörigen Klassen zuordnen.⁶⁶⁷ Als Klassifizierungskriterium eignet sich der Nutzungsgrad der Informationen.⁶⁶⁸ Eine Klasse bezeichnet die geeignete Speicherebene⁶⁶⁹, auf welche die Informationen verlagert werden sollen. Damit jede Information auf die jeweils sinnvollste Speicherebene verlagert werden kann, sollte jede Information in die automatisierte Klassifikation einbezogen werden.⁶⁷⁰

5.1.3.2.3 Berücksichtigung des Anwender- und Administratorenwissens

Trotz der Automatisierung sollte das Wissen der Anwender und Administratoren über die Systeme und Informationen berücksichtigt werden.⁶⁷¹ Anwender und Administratoren können wichtige Hinweise zur Verwendung der Informationen geben, beispielsweise auf Basis der aktuellen und zukünftigen Projekt- bzw. Auftragslage oder der Verwendung von Informationen in verschiedenen Abteilungen eines Unternehmens.⁶⁷² Insbesondere bei der Definition der Klassifizierungsfunktionen und bei der Festlegung, ab welchem Nutzungsgrad eine Information einer bestimmten Klasse angehören soll, ist dies von Relevanz. Außerdem ist das Wissen der Anwender und Administratoren für die Validierung der Klassifikation wichtig, da diese die Besonderheiten der Systeme und

⁶⁶⁶ Vgl. Chen / Information valuation / 135-146

⁶⁶⁷ Vgl. Chen / Information valuation / 135-146; Mesnier et al. / File classification / 44-51; Shah et al. / Classification for ILM / 1-15; Verma et al. / Very Large File Systems / 160-168; Vgl. Kapitel 2.4.2

⁶⁶⁸ Vgl. Kapitel 2.4.1

⁶⁶⁹ Vgl. Kapitel 5.1.2.3

⁶⁷⁰ Vgl. Bhagwan et al. / Management of Data Storage / ; Shah et al. / Classification for ILM / 12-14; Turczyk / Wertzuweisung / 462

⁶⁷¹ Vgl. hierzu Bhagwan et al. / Management of Data Storage /; Douglass et al. / Delete-Optimized Storage /; Patrascu et al. / Autonomic computing / . Informationen bzw. Dokumente in einem Dokumentenmanagementsystem können inhaltlich zusammengehörig sein, weil sie bspw. von einer Abteilung erstellt und genutzt werden oder weil sie einem bestimmten Projekt zugeordnet sind. Eine bloße Verlagerung aufgrund des Nutzungsgrades wäre in diesem Fall nicht sinnvoll. Hier müssen die Anwender der Informationen entsprechende Hinweise zur Verwaltung der Informationen geben.

⁶⁷² Vgl. Short / ILM / 2-37; Gillet, Mendel / Organic IT / 483-502; Marcial, Matthes, Hartmann / Digitale Archivierung / 74-81

Informationen am besten kennen.⁶⁷³ Konzepte zur automatisierten Informationsklassifizierung sollten deshalb die Möglichkeit bieten, die Klassenzugehörigkeit auch nach der automatischen Klassifizierung manuell anzupassen. Dies würde es ermöglichen, nicht sinnvoll erscheinende Klassenzuordnungen noch vor der Informationsverlagerung zu revidieren, was insbesondere bei automatisierten Vorschlägen zur Löschung von Informationen wichtig ist.

5.1.3.2.4 Berücksichtigung rechtlicher Rahmenbedingungen

Neben dem Anwender- und Administratorenwissen sollte auch die Berücksichtigung rechtlicher Rahmenbedingungen⁶⁷⁴ ein Bestandteil der Validierung der Klassenzugehörigkeit sein, womit eine weitere wichtige Anforderung an Konzepte zur automatisierten Klassifizierung von Informationen identifiziert werden kann.⁶⁷⁵ Es sollten Experten hinzugezogen werden, die die rechtlichen Anforderungen kennen. Die rechtlichen Anforderungen sind sehr komplex und werden vom Gesetzgeber häufig verändert bzw. angepasst.⁶⁷⁶ Muss eine Information aufgrund rechtlicher Restriktionen aufbewahrt werden, so darf sie nicht gelöscht werden, obwohl eventuell für einen längeren Zeitraum nicht auf sie zugegriffen wurde. Andererseits existieren Informationen, die nach einem gesetzlich vorgeschriebenen Zeitraum gelöscht werden müssen, obwohl oft darauf zugegriffen worden ist. In beiden Fällen ist eine nachträgliche Anpassung der Klassenzugehörigkeit der betreffenden Information nach einer automatisierten und auf Grundlage des Nutzungsgrads vorgenommenen Klassifikation notwendig.

5.1.3.2.5 Kostenreduktion und Wirtschaftlichkeit

Das Datenwachstum in Unternehmen führt zu hohen Speicher- und Administrationskosten.⁶⁷⁷ Die Kosten für die Administration der Datenträger bzw. Datenspeicher, wie beispielsweise die Datensicherung und Spiegelung der Daten sowie die sichere Aufbewahrung sind um das vier- bis achtfache höher als die Beschaffungskosten der Datenträ-

⁶⁷³ Vgl. Keeney, Raiffa / Multiple Objectives /

⁶⁷⁴ Vgl. Kapitel 2.4.3, Kapitel 3.4.2

⁶⁷⁵ Vgl. Mont, Beato / Obligation Policies / 51-55; Short / ILM / 2-37; Lotlikar, Mohania / Policies ILM / 612-621

⁶⁷⁶ Vgl. Johnson, Agrawal / Law and Technology / ; Kaarst-Brown, Kelly / IT-Governance /

⁶⁷⁷ Vgl. Abd-El-Malek et al. / self-storage / 55-62; Allen / waste storage dollars / ; Thome, Sollbach / Modelle des ILM / 22-30

ger.⁶⁷⁸ Wesentliches Ziel von ILM ist die Kostensenkung. Die automatisierte Informationsklassifizierung sollte daher nicht zuletzt zu einer Reduktion der Administrationskosten für Datenspeicher⁶⁷⁹ führen. Es muss eine Reduktion der Aufbewahrungskosten aufgrund der Speicherung der Informationen in der jeweils angemessenen Speicherebene erfolgen,⁶⁸⁰ damit von einem effektiven ILM gesprochen werden kann. Darüber hinaus muss die Leistungsfähigkeit der Systeme berücksichtigt werden. Weiterhin muss der Aufwand für die Implementierung und Anwendung des ILM-Konzepts in einem angemessenen Kosten- und Nutzenverhältnis stehen.⁶⁸¹

5.1.3.2.6 Berücksichtigung großer Datenmengen

Der Einsatz von ILM ist nur bei großen und/oder schnell wachsenden Datenbeständen⁶⁸² wirtschaftlich sinnvoll.⁶⁸³ Bei kleinen Datenbeständen steht der Aufwand für die Klassifizierung der Daten in keinem angemessenen Verhältnis zur Kostensenkung.⁶⁸⁴ Aus diesem Grund sollte ein Konzept zur automatisierten Klassifizierung die Klassifizierung großer Datenmengen berücksichtigen⁶⁸⁵ und es sollte überprüft werden, in welchen Zeitabständen Informationen zwischen den Speicherebenen verlagert werden müssen.⁶⁸⁶ Gleichzeitig sollte im Sinne der Kostensenkung die Leistungsfähigkeit der Systeme berücksichtigt werden.⁶⁸⁷

⁶⁷⁸ Vgl. Glazer / Measuring Information / 99-110; Kaplan, Roy, Srinivasaraghavan / Demand data storage / 1

⁶⁷⁹ Vgl. Mesnier et al. / File classification / 44-51

⁶⁸⁰ Vgl. Shah et al. / Classification for ILM / 1-15

⁶⁸¹ Vgl. Kapitel 5.2.1

⁶⁸² Mit einem physischen Volumen ab einem Terabyte. Vgl. hierzu Kapitel 1.1.

⁶⁸³ Zur Abschätzung des zukünftigen Datenwachstums kann eine Prognose durchgeführt werden. Vgl. hierzu Kapitel 5.1.2.2

⁶⁸⁴ Vgl. Heinrich, Stelzer / Informationsmanagement / 275

⁶⁸⁵ Vgl. Bhagwan et al. / Management of Data Storage / ; Shah et al. / Classification for ILM / 1-15

⁶⁸⁶ Vgl. Kapitel 5.1.2

⁶⁸⁷ Vgl. Kapitel 1.3

5.2 Konstruktion und Implementierung eines Verfahrens zur Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM

5.2.1 Anforderungen an die Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM

Da in dieser Arbeit die Wirtschaftlichkeit der automatisierten Klassifizierung von Daten untersucht wird, wurden in Kapitel 5.1.3.2 Anforderungen an Konzepte zur automatisierten Klassifizierung unstrukturierter Daten erarbeitet. Als Anforderungen wurden die „Verwendung des Nutzungsgrads als Klassifizierungskriterium“, die „Implementierung geeigneter Klassifizierungsfunktionen“, die „Berücksichtigung des Anwender- und Administratorenwissens“, die „Berücksichtigung rechtlicher Rahmenbedingungen“, die „Kostenreduktion und Wirtschaftlichkeit“ und die „Berücksichtigung großer Datenmengen“ definiert. Neben den Ergebnissen der Literaturrecherche⁶⁸⁸ fließen diese Anforderungen mit in die Anforderungen an die Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten⁶⁸⁹ im ILM ein:⁶⁹⁰

- Analyse der Kosten- und Nutzenstruktur: Um die Wirtschaftlichkeit des ILM beurteilen zu können, muss die Kosten- und Nutzenstruktur des ILM analysiert werden.⁶⁹¹ Anschaffungs- und Betriebskosten, die für die Implementierung von Konzepten zur Klassifizierung von Informationen anhand des Nutzungsgrads⁶⁹², für die Klassifizierung⁶⁹³ und Verwaltung der Informationen in den unterschiedlichen Speicherebenen und zur Erfüllung rechtlicher Rahmenbedingungen⁶⁹⁴ anfallen, müssen in die Betrachtungen einfließen. Relevante Kosten sind bspw. Be-

⁶⁸⁸ Vgl. Kapitel 2.4, Kapitel 3.4, Kapitel 4 und Kapitel 5.1.3.2

⁶⁸⁹ Inwieweit die Anforderungen durch die in dieser Arbeit demonstrierten bzw. durchgeführten Wirtschaftlichkeitsanalyse erfüllt werden, wird in den Kapiteln 5.3 und 5.4 erörtert. Hierzu wird ein Merkmalsvergleich zwischen definierten Anforderungen und erzielten Ergebnissen vorgenommen. Vgl. hierzu Peffers et al. / Research Methodology / 55-56; Heinrich / Evaluationsforschung / 9; Kapitel 1.3.

⁶⁹⁰ Technische Anforderungen, wie Antwortzeiten, Schnittstellen oder Verfügbarkeit werden nicht betrachtet, da es nicht Ziel der Arbeit ist, eine Softwareanwendung zu entwickeln. Ziel der prototypischen Implementierung zur Wirtschaftlichkeitsanalyse ist es, die Wirtschaftlichkeit der automatisierten Klassifizierung im ILM zu untersuchen. Vgl. zur Definition und Abgrenzung von Anforderungen in Bezug auf das Artefakt und das Forschungsziel Peffers et al. / Research Methodology / 56.

⁶⁹¹ Vgl. Kapitel 4.1

⁶⁹² Vgl. hierzu die Anforderung zur „Verwendung des Nutzungsgrads als Klassifizierungskriterium“, Kapitel 5.1.3.2.1

⁶⁹³ Vgl. hierzu die Anforderung zur „Implementierung geeigneter Klassifizierungsfunktionen“, Kapitel 5.1.3.2.2

⁶⁹⁴ Vgl. hierzu die Anforderung zur „Berücksichtigung rechtlicher Rahmenbedingungen“, Kapitel 5.1.3.2.4; Kapitel 2.4.3; Kapitel 3.4.2

ratungs- und Implementierungskosten, Speicher- und Administrationskosten, Kosten für die Erhebung der Nutzungsgrade unstrukturierter Daten, Schulungskosten, Verlagerungs- bzw. Migrationskosten, Kosten für die Datenübertragung und Kosten, die bei der Nutzung von Public-Cloud-Diensten⁶⁹⁵ anfallen.

- Analyse der Beziehungszusammenhänge von Kosten und Nutzen: Die Beziehungszusammenhänge zwischen den einzelnen Kosten- und Nutzenarten müssen analysiert werden.⁶⁹⁶ Es muss untersucht werden, welche Kosten welchen Nutzen erzeugen oder umgekehrt, welche Nutzenart welche Kosten verursacht.⁶⁹⁷ Wesentliches Ziel von ILM ist die kostengünstige Verwaltung von Informationen. ILM sollte zu einer Senkung der Kosten, die für die Verwaltung der Informationen anfallen, führen. Hierfür eignen sich bspw. die Untersuchung und der Vergleich der Wirtschaftlichkeit der Verwaltung der Informationen mit ILM und ohne ILM. Dazu können die Methoden der Wirtschaftlichkeitsrechnung angewendet werden, wie bspw. die Kostenvergleichsrechnung.⁶⁹⁸
- Automatisierung der Wirtschaftlichkeitsanalyse: Die Analyse der Beziehungszusammenhänge und der Wirtschaftlichkeit der Klassifizierung sollte zumindest teilweise automatisiert werden, da die Kosten für die Verwaltung und Klassifizierung von Informationen entsprechend der Zugriffe bzw. des Nutzungsgrads⁶⁹⁹ variieren. Hierfür eignet sich die Implementierung von Kosten- und Verlagerungsfunktionen, die die Kosten für die Klassifizierung und Verwaltung der Informationen entsprechend der verwendeten Speicherebene berechnen.⁷⁰⁰ Weiterhin eignet sich die Anwendung der Simulationstechnik⁷⁰¹ und der Kostenver-

⁶⁹⁵ Vgl. Kapitel 3.4

⁶⁹⁶ Vgl. hierzu Kapitel 4.1

⁶⁹⁷ Vgl. hierzu die Anforderung zur „Kostenreduktion und Wirtschaftlichkeit“, Kapitel 5.1.3.2.5.

⁶⁹⁸ Vgl. Kapitel 4.1.

⁶⁹⁹ Die Zugriffe auf die Informationen erfolgen zufällig und sind schwer vorhersehbar.

⁷⁰⁰ Durch die teilweise Automatisierung der Wirtschaftlichkeitsanalyse können Planungsaufwände bzw. Kosten, die für die Durchführung der Wirtschaftlichkeitsanalyse anfallen, reduziert werden. Vgl. hierzu Becker / Wirtschaftswissen / 180.

⁷⁰¹ Vgl. hierzu Kapitel 4.2.2.2. Das Simulationsmodell sollte ohne großen Aufwand flexibel anpassbar sein, um die Parameter der Simulation, wie Intervalle für Klassenzugehörigkeiten oder Nutzungsgrade der Informationen, modifizieren zu können.

gleichsrechnung⁷⁰². Der Planungsaufwand der statischen Methode der Kostenvergleichsrechnung ist im Vergleich zu dynamischen Methoden der Wirtschaftlichkeitsrechnung gering.⁷⁰³ So kann sichergestellt werden, dass Aufwand und Nutzen der durchgeführten Wirtschaftlichkeitsanalyse im ILM in einem angemessenen Verhältnis zueinander stehen.⁷⁰⁴ Es können verschiedene Szenarien⁷⁰⁵ mit unterschiedlichen Nutzungsgraden, Klassifizierungsfunktionen, Speicher-ebenen und unterschiedlich großen Datenmengen⁷⁰⁶ simuliert werden, um die jeweiligen Klassifizierungs- und Verwaltungskosten der Informationen zu ermitteln.⁷⁰⁷

- Berücksichtigung des Cloud Computings zur Verwaltung der Informationen: Zusätzlich zu den unternehmensinternen Speichermedien sollte das Cloud Computing zur wirtschaftlichen Verwaltung der Informationen betrachtet werden.⁷⁰⁸
- Untersuchung der Ergebnisse: Die Ergebnisse der Wirtschaftlichkeitsanalyse müssen untersucht und ggf. ergänzt werden, um die Wirtschaftlichkeit von ILM beurteilen zu können.⁷⁰⁹
- Berücksichtigung des Anwender- und Administratorenwissens: Trotz der Notwendigkeit der Automatisierung sollten Anwender und Administratoren bei der Klassifizierung der Informationen⁷¹⁰ und Auswertung der Ergebnisse der Wirtschaftlichkeitsanalyse eingebunden werden. Beispielsweise kann aufgrund der Ergebnisse der Wirtschaftlichkeitsanalyse von Daten, auf die keine Zugriffe erfolgt sind, die Löschung der Daten aus Kostengesichtspunkten als zweckmäßig

⁷⁰² Vgl. Kapitel 4.2.1.1

⁷⁰³ Vgl. Kapitel 4.2.1.2; Becker / Wirtschaftswissen / 180

⁷⁰⁴ Vgl. Kapitel 4.2.1, Kapitel 5.1.3.2.5, Kapitel 5.1.3.2.6, Kapitel 5.3

⁷⁰⁵ Vgl. Kapitel 5.2.3.1.1.1

⁷⁰⁶ Vgl. hierzu die Anforderung zur „Berücksichtigung großer Datenmengen“, Kapitel 5.1.3.2.6

⁷⁰⁷ Es gelten entsprechend die Anforderungen an Konzepte zur automatisierten Klassifizierung unstrukturierter Daten. Vgl. hierzu Kapitel 5.1.3.2.

⁷⁰⁸ Vgl. Kapitel 5.2.3.1.1.1.2

⁷⁰⁹ Vgl. Kapitel 4.1; Kapitel 5.2.3.3.3; Kapitel 5.3; Kapitel 5.4

⁷¹⁰ Vgl. hierzu die Anforderung zur „Berücksichtigung des Anwender- und Administratorenwissens“, Kapitel 5.1.3.2.3. Es muss vermieden werden, dass zusammengehörige Daten, die bspw. einem Geschäftsprozess angehören, durch Konzepte zur automatisierten Klassifizierung auf unterschiedlichen Speicherebenen abgelegt werden. Vgl. hierzu Kapitel 5.2.2.1.3.1.

erscheinen. Wenn Anwendern und Administratoren allerdings bekannt ist, dass diese Daten in naher Zukunft für ein bestimmtes Projekt benötigt werden, müssen diese jedoch aufbewahrt werden.

5.2.2 Konstruktion eines Verfahrens zur Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM

Das in Kapitel 5.1.2 konstruierte ILM-Vorgehensmodell eignet sich für die Verwaltung strukturierter und unstrukturierter Daten.⁷¹¹ In der vorliegenden Arbeit wird die Verwaltung unstrukturierter Daten betrachtet, da insbesondere unstrukturierte Daten zum Gegenstandsbereich von ILM zählen.⁷¹² Aufbauend auf die in Abbildung 5-7 hervorgehobene Phase der „Klassifizierung der Daten“ wird ein Verfahren zur Wirtschaftlichkeitsanalyse im ILM konstruiert und anschließend prototypisch implementiert.⁷¹³

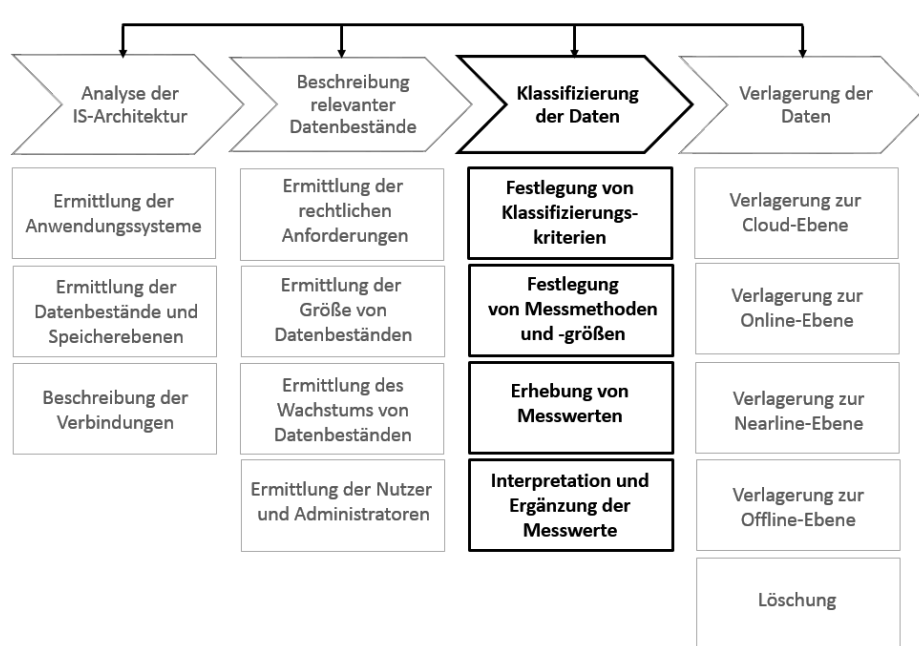


Abb. 5-7: Klassifizierung der Daten

⁷¹¹ Vgl. hierzu zur automatisierten Klassifizierung von strukturierten Daten in einem Data-Warehouse-System Kosler, Matthesius, Stelzer / Klassifizierung von Informationen / 129-145.

⁷¹² Vgl. Kaiser, Smolnik, Riempp / ILM-Framework / 484; Kanakamedala, Kaplan, Srinivasaraghavan / Data storage / 1-4; Etwa achtzig Prozent der Daten eines Unternehmens sind unstrukturierte Daten (Vgl. Gantz et al. / Digital Universe / 12; Schoop / Informationsmanagement / 556-568; Vouk / Cloud Computing / 242-243; McKinsey / Challenge Big Data / 2-3. Etwa vierundneunzig Prozent des weltweiten Datenaufkommens sind unstrukturierte Daten (Vgl. Gantz et al. / Exploding Universe / 2-5; Gantz, Reinsel / Extracting Value / 1-2).

⁷¹³ Vgl. Kapitel 1.3. Die wesentliche Teilaufgabe von ILM ist die Klassifizierung der Daten. Vgl. Kosler, Matthesius, Stelzer / Klassifizierung von Informationen / 129-145.

Es wird erörtert, wie bei der „Klassifizierung der Daten“ vorzugehen ist.⁷¹⁴ Es werden Daten zur Abbildung verschiedener Lebenszyklen⁷¹⁵ generiert sowie ein Simulationsmodell⁷¹⁶ zur Analyse der Wirtschaftlichkeit implementiert.⁷¹⁷ Dies ist neben den Aspekten der Integration des Cloud Computings in das ILM und der Betrachtung der relevanten Kosten- und Nutzenverhältnisse ein weiteres wesentliches Unterscheidungsmerkmal der vorliegenden Arbeit im Vergleich zu publizierten Forschungsergebnissen. Hierbei wurden lediglich ausgewählte Daten in einzelnen Organisationen betrachtet.⁷¹⁸ Welchen Einfluss verschiedene Lebenszyklen auf die Verwaltung und Klassifizierung der Daten haben, und welche Auswirkungen auf die Kosten entstehen, wurde in diesen Untersuchungen nicht betrachtet. Es kommen Klassifizierungs-, Kosten- und Verlage-
rungsfunktionen⁷¹⁹ zur Anwendung. Die Demonstration und Evaluation der Wirtschaftlichkeitsanalyse erfolgen anschließend in den Kapiteln 5.3 und 5.4. Im Rahmen der prototypischen Implementierung werden die Daten nicht physisch auf unterschiedliche Speicherebenen verlagert. Allerdings wird in Kapitel 5.3 eine Kosten- und Nutzenanalyse bezüglich der Verwaltung der Daten für unterschiedliche Lebenszyklen durchgeführt. Daraus wird ersichtlich, mit welchem Kosten- und Nutzenverhältnis bei der Verlagerung der Daten auf die Speicherebenen zu rechnen ist. Hieraus werden Handlungsempfehlungen abgeleitet, wie bezogen auf die Art des Lebenszyklus bzw. des Nutzungsgrads der Informationen und der Datenmenge eine wirtschaftlich sinnvolle Verwaltung der Daten erzielt werden kann.⁷²⁰

⁷¹⁴ Vgl. Kapitel 5.2.2.1.1, 5.2.2.1.2 und 5.2.2. Neben der Beschreibung der Vorgehensweise und den zu erwartenden Ergebnissen werden Hinweise zu Software-Werkzeugen gegeben, die zur Umsetzung der Teilphasen unterstützend eingesetzt werden können, um beispielsweise die Metadaten unstrukturierter Daten zu erfassen.

⁷¹⁵ Vgl. Kapitel 5.2.2.1.3

⁷¹⁶ Vgl. Greasley / Simulation Modelling / 2-5; Banks / Simulation / 3-9

⁷¹⁷ Vgl. Kapitel 5.2.3

⁷¹⁸ Vgl. Turczyk / Organisation ILM / 110; Douglass et al. / Delete-Optimized Storage / ; Chen / Information valuation / 135-146; Turczyk / Wertzuweisung / 467

⁷¹⁹ Vgl. Kapitel 5.2.2.1.3

⁷²⁰ Vgl. Kapitel 5.3.

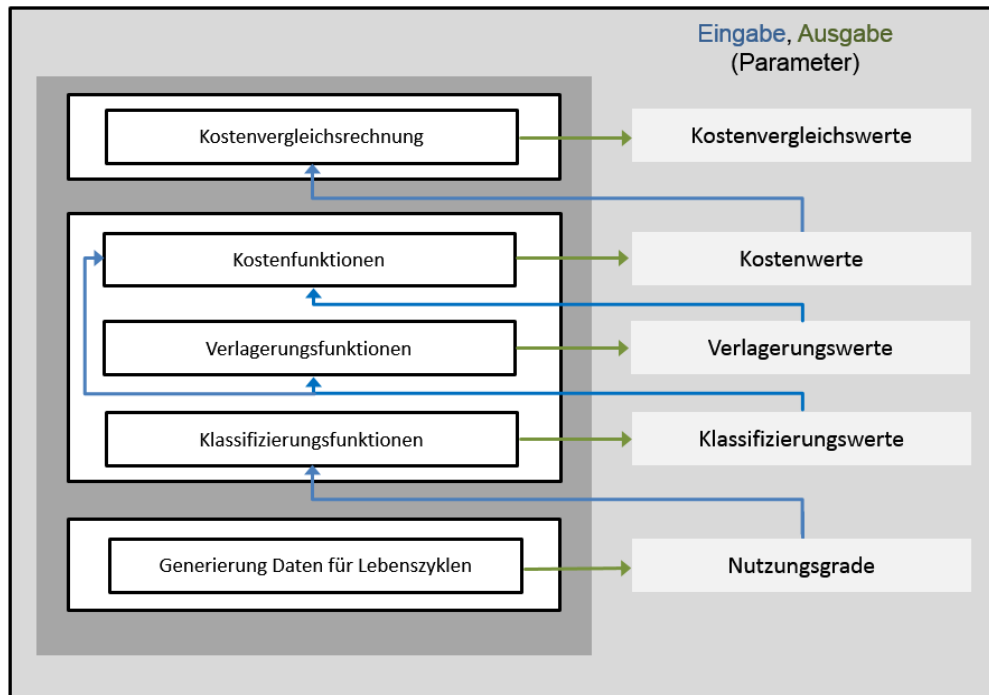


Abb. 5-8: Aufbau und Ablauf der Wirtschaftlichkeitsanalyse

In Abbildung 5-8 sind Aufbau und Ablauf der Wirtschaftlichkeitsanalyse dargestellt. Die generierten Nutzungsgrade⁷²¹ werden an die Klassifizierungsfunktionen⁷²² übergeben.⁷²³ Die Klassifizierungsfunktionen berechnen die Klassenwerte auf Basis der Nutzungsgrade.⁷²⁴ Die berechneten Klassenwerte fließen in die Berechnung der Verlagerungswerte⁷²⁵ und Kostenwerte⁷²⁶ ein.⁷²⁷ Die Verlagerungs- und Kostenwerte werden durch die entsprechenden Verlagerungs- und Kostenfunktionen ermittelt. Die Kostenwerte, die die Kosten für die automatisierte Klassifizierung und die Verlagerungsvorgänge enthalten, werden an die abschließende Kostenvergleichsrechnung⁷²⁸ übergeben, um die Wirtschaftlichkeit der automatisierten Klassifizierung im ILM zu untersu-

⁷²¹ Vgl. Kapitel 5.2.2.1.3.1

⁷²² Vgl. Kapitel 5.2.3.1

⁷²³ Vgl. zur Festlegung der Klassifizierungskriterien, Messmethoden und -größen sowie der Erhebung der Messwerte im Rahmen der Wirtschaftlichkeitsanalyse Kapitel 5.2.2.1.1, Kapitel 5.2.2.1.2 und Kapitel 5.2.2.1.3.

⁷²⁴ Vgl. Kapitel 5.2.3.1.1

⁷²⁵ Vgl. Kapitel 5.2.3.2

⁷²⁶ Vgl. Kapitel 5.2.3.3

⁷²⁷ Vgl. Kapitel 5.3

⁷²⁸ Vgl. Kapitel 4.2.1.1

chen.⁷²⁹ Hierbei wird untersucht, welcher monetär messbare Nutzen durch die Anwendung von ILM erzielt werden kann. Wirtschaftlichkeit⁷³⁰ von ILM ist gegeben, wenn die Verwaltung der Daten im ILM zu Kostensenkungen führt.⁷³¹

5.2.2.1 Klassifizierung der Daten

Die aus den Phasen „Analyse der IT-Architektur“ und „Beschreibung relevanter Datenbestände“ resultierenden Ergebnisse werden als gegeben betrachtet⁷³² und im Folgenden vorgestellt. Hinsichtlich der eingesetzten Anwendungssysteme⁷³³ und Speicherebenen wird die Annahme getroffen, dass die Daten in Dateiordnern⁷³⁴ verwaltet werden.⁷³⁵ Für die Verwaltung der Daten werden, in Abhängigkeit vom untersuchten Szenario,⁷³⁶ die Speicherebenen „Cloud“, „Online“, „Nearline“ und „Archiv“ verwendet. Bezüglich der Speicherebene „Cloud“ erfolgt die Nutzung des Dienstes Infrastructure-as-a-Service (IaaS).⁷³⁷ Aufgrund rechtlicher Anforderungen⁷³⁸ wird angenommen, dass die Daten während des betrachteten Zeitraums von fünf Jahren, nicht gelöscht werden.⁷³⁹ Die Untersuchungen für die verschiedenen Nutzungsgrade der Informationen erfolgen mit ver-

⁷²⁹ Vgl. zur Interpretation und Ergänzung der Messwerte Kapitel 5.2.3.3.3. Vgl. zur Demonstration und Evaluation der Wirtschaftlichkeitsanalyse Kapitel 5.3 und Kapitel 5.4.

⁷³⁰ Vgl. Kapitel 4.1

⁷³¹ Vgl. Kapitel 5.3

⁷³² Zur Beschreibung und Anwendung der Phasen „Analyse der IT-Architektur“ und „Beschreibung relevanter Datenbestände“ vgl. Kapitel 5.1.2.1 und 5.1.2.2.

⁷³³ Es wird in dieser Arbeit nicht der Lebenszyklus von Anwendungssystemen betrachtet, sondern der Lebenszyklus der in den Systemen verwalteten Daten. Die Betrachtung des Lebenszyklus von Anwendungssystemen erfolgt im Rahmen des Konzepts des Application Lifecycle Managements (Vgl. hierzu Kaariainen, Valimäki / Application Lifecycle Management / 55-65).

⁷³⁴ Beispielsweise in Dateiordnern des Betriebssystems Microsoft Windows. Denkbar wäre auch die Verwaltung der Daten mit Hilfe eines Dokumentenmanagementsystems (DMS). Ein Beispiel für ein DMS ist Microsoft Sharepoint. Vgl. hierzu Laahs, McKenna, Vanamo / Sharepoint-Technologien / 71-100.

⁷³⁵ Vgl. Kapitel 5.1.2.1

⁷³⁶ Vgl. Kapitel 5.2.3.1; Es werden zwei Szenarien zur Klassifizierung der Daten betrachtet. Im ersten Szenario wird angenommen, dass die Speicherebene „Cloud“ aufgrund rechtlicher Anforderungen (Vgl. Kapitel 3.4.2) nicht verwendet wird (Vgl. Kapitel 5.2.2.1.3, Kapitel 5.2.3.3.3). Im zweiten Szenario wird untersucht, inwiefern die zusätzliche Nutzung von Cloud-Speicher bzw. IaaS-Diensten, Auswirkungen auf das Kosten- und Nutzenverhältnis der Verwaltung unstrukturierter Daten hat. Vgl. hierzu Kapitel 5.2.3.1.1.1.2.

⁷³⁷ Vgl. Kapitel 3.2.2; Vgl. Kapitel 5.2.3.3.2.1.2 für den verwendeten Cloud-Anbieter.

⁷³⁸ Vgl. Kapitel 2.4.3, Kapitel 3.4.2

⁷³⁹ Vgl. Kapitel 5.2.2.1.3, Kapitel 5.2.3.3.3 und Kapitel 5.3.

schiedenen Datenmengen.⁷⁴⁰ Es wird angenommen, dass das jährliche Datenwachstum für den betrachteten Zeitraum von fünf Jahren, dreißig Prozent je betrachteter Datenmenge beträgt.⁷⁴¹

5.2.2.1.1 Festlegung von Klassifizierungskriterien

Im Rahmen der prototypischen Implementierung⁷⁴² wird der Nutzungsgrad als Kriterium für die Klassifizierung von Informationen herangezogen.⁷⁴³ Das Klassifizierungskriterium Nutzungsgrad lässt sich ohne großen Aufwand ermitteln und automatisiert messen.⁷⁴⁴ Durch die Verwendung des Nutzungsgrads und der jeweiligen Datenmenge⁷⁴⁵ wird die Klassifizierung entsprechend vereinfacht. Idealerweise sollte bei der Klassifizierung das Anwender- und Administratorenwissen einbezogen werden oder rechtliche Anforderungen berücksichtigt werden.⁷⁴⁶ Ein Zugriff erfolgt beispielsweise beim Erstellen und Editieren der Daten.⁷⁴⁷ Die Zugriffshäufigkeit variiert im Zeitverlauf.⁷⁴⁸ Es wird ermittelt, zu welchen Zeitpunkten auf eine Information zugegriffen wurde. So kann bestimmt werden, ob Informationen genutzt werden oder ob für einen längeren Zeitraum nicht darauf zugegriffen wurde.⁷⁴⁹ Die Klassifizierungskriterien und die Lebenszyklen der Daten werden auf Basis von bereits vorhandenen Forschungsergebnissen zu Verhaltensstudien zur Nutzung von Informationen erzeugt.⁷⁵⁰

⁷⁴⁰ Vgl. Kapitel 5.3.

⁷⁴¹ Vgl. Kapitel 5.3. zur Analyse der Wirtschaftlichkeit unter Berücksichtigung verschiedener Datenmengen. Für das jeweils erste Jahr werden Datenmengen von 1 Terabyte, 10 Terabyte und 100 Terabyte betrachtet. Vgl. hierzu Anhang 1.

⁷⁴² Vgl. Kapitel 5.2.3

⁷⁴³ Vgl. Kapitel 2.4.1

⁷⁴⁴ Vgl. Bhagwan et al. / Management of Data Storage /; Chen / Information valuation / 135-146; Chandra, Gehani, Yu / Automated Storage / 12; Douglass et al. / Delete-Optimized Storage /; Gavish, Sheng / Dynamic file migration / 177-189; Shah et al. / Classification for ILM / 1-15; Skyrme / Add value to business / 20-25

⁷⁴⁵ Vgl. Kapitel 5.2.2.1

⁷⁴⁶ Vgl. Kapitel 5.1.3.2

⁷⁴⁷ Vgl. Herrmann, Spruth / Einführung in z/OS / 90; Gibson, Miller, Long / File-Usage / 641; Zadok et al. / Reducing Storage Management Costs / 101-102

⁷⁴⁸ Vgl. Bhagwan et al. / Management of Data Storage /; Chen / Information valuation / 135-146; Shah et al. / Classification for ILM / 1-15

⁷⁴⁹ Vgl. Chen / Information valuation / 135-146

⁷⁵⁰ Vgl. Kapitel 5.2.2.1.3

5.2.2.1.2 Festlegung von Messmethoden und -größen

Die Klassifizierung der Informationen wird automatisiert durchgeführt.⁷⁵¹ Es werden Funktionen⁷⁵² implementiert, die die Informationen, basierend auf den Klassifizierungskriterien, klassifizieren und den zugehörigen Klassen zuordnen.⁷⁵³ Im Rahmen der prototypischen Implementierung werden die Zugriffe im Zeitverlauf bzw. der Nutzungsgrad basierend auf den vorhandenen Forschungsergebnissen mit Hilfe des Softwarewerkzeugs Microsoft Excel 2013 erzeugt.⁷⁵⁴ Somit kann die Klassifizierung der Daten, unabhängig von bestimmten Gegebenheiten einer Unternehmung oder Organisation, für verschiedene Nutzungsgrade, umgesetzt werden.⁷⁵⁵ Die erzeugten Daten repräsentieren typische Nutzungsgrade von unstrukturierten Daten, wie PDF-Dateien oder Microsoft-Office-Dateien.⁷⁵⁶ Eine Klasse bezeichnet die jeweilige Speicherebene,⁷⁵⁷ auf welche die Informationen verlagert⁷⁵⁸ werden können. Im Anschluss an die automatisierte Klassifizierung der Informationen wird die Anzahl der Verlagerungs- bzw. Migrationsvorgänge für die einzelnen Klassifizierungsläufe ermittelt.⁷⁵⁹ Die abschließende Interpretation bzw. Auswertung der Messergebnisse und Verlagerungsvorgänge liefert wichtige Hinweise zur Wirtschaftlichkeit der durchgeführten Klassifizierung.⁷⁶⁰

⁷⁵¹ Vgl. Konzepte zur automatisierten Klassifizierung Kapitel 5.1.3

⁷⁵² Die Klassifizierungsfunktionen und das Simulationsmodell werden in Microsoft Excel 2013 erstellt (Vgl. hierzu Kapitel 5.2.2.1.3). Microsoft Excel eignet sich für den Aufbau von Simulationsmodellen und für die Durchführung von Experimenten. Vgl. hierzu Allman, Laurito, Loh / Simulation Modeling in Excel / 6-11; Guerrero / Modeling and Simulation / 3-5.

⁷⁵³ Es wird angenommen, dass die Klassifizierung der Daten einmal im Monat durchgeführt wird. Dabei werden die Nutzungsgrade der Informationen des jeweiligen Vormonats in die Klassifizierung einbezogen. Vgl. Kapitel 5.2.3.1.

⁷⁵⁴ Für die Ermittlung des Nutzungsgrads von unstrukturierten Daten können ebenfalls Softwarewerkzeuge, wie bspw. das „Metadata Extraction Tool“ (Vgl. NLNZ / Metadata Extraction Tool /), verwendet werden. Hierzu können die relevanten Informationen über die Zugriffe auf die Daten genutzt werden, die in Form von XML-basierten Daten innerhalb der Dateien bzw. Dokumente systemseitig gespeichert werden (Vgl. Wootton / Developing Quality Metadata / 200-201; Witten et al. / Documents and Metadata / 389-405; Baker / Adobe Acrobat / 194-195; Antonovich / Office User's Guide / 222-231).

⁷⁵⁵ Vgl. Kapitel 5.2.2.1.3

⁷⁵⁶ Vgl. Kapitel 5.2.2.1.3

⁷⁵⁷ Vgl. Kapitel 5.1.2.3

⁷⁵⁸ Vgl. Kapitel 5.1.2.4

⁷⁵⁹ Vgl. Kapitel 5.2.3.3.3; Hierzu wird ebenfalls eine Funktion in Microsoft Excel erzeugt (Vgl. Kapitel 5.2.3.2).

⁷⁶⁰ Vgl. Kapitel 5.2.3.3.3, Kapitel 5.3.

5.2.2.1.3 Erhebung von Messwerten

5.2.2.1.3.1 Lebenszyklen der Daten

Es existieren Forschungsergebnisse zu Verhaltensstudien zur Nutzung von Informationen. Es werden sechs typische Lebenszyklen von Daten herausgearbeitet, die am häufigsten in den Forschungsprojekten identifiziert wurden.⁷⁶¹ Der am häufigsten vorkommende Lebenszyklus beschreibt einen rapiden Rückgang der Zugriffshäufigkeiten, kurz nach der Entstehung der Daten. Die weiteren Lebenszyklen beschreiben jeweils steigende, konstante, saisonale und trend-saisonale Lebenszyklen bzw. Nutzungsgrade. Die in diesen Studien untersuchten Informationen sind beispielsweise unstrukturierte Daten aus Forschung und Entwicklung, E-Mails, Finanz- und Medizindaten aus verschiedenen Forschungseinrichtungen und Unternehmen unterschiedlicher Branchen. In Anlehnung an diese Lebenszyklen werden die Nutzungsgrade bzw. Lebenszyklen generiert bzw. erzeugt.⁷⁶² Ein Lebenszyklus bildet den Nutzungsgrad von Daten ab,⁷⁶³ die sich in einem Dateiordner befinden oder zu einem bestimmten Geschäftsprozess gehören.⁷⁶⁴ Da Nutzungsgrade unstrukturierter Daten in der Realität schwer planbar sind und zufällig

⁷⁶¹ Vgl. Agrawal et al. / Study File-System Metadata / 31-45; Bennett, Bauer, Kinchlea / NFS environments / 33-40; Chen / Information valuation / 141; Douceur, Bolosky / Study of File-System / 64-66; Gibson, Miller, Long / File-Usage / 639-648; Jensen, Reed / Activity Computer Environment / 387-396; Miller, Katz / Behavior Supercomputer / 51-55; Lundberg / Lifecycle Management / 5; Merrill, Thanhardt / Mass Storage / 117-121; Chandra, Gehani, Yu / Automated Storage / 5-12; Miller, Katz / Analysis File Migration / 421-434; Moore / ILM / ; Satyanarayanan / Study Lifetimes / 96-108; Schmitz / Migrationsstrategie / 21-24; Shah et al. / Classification for ILM / 1-15; Spasojevic, Satyanarayanan / Study File System / 200-222; Zadok et al. / Reducing Storage Management Costs / 101-105

⁷⁶² Vgl. hierzu Kapitel 5.2.2.1.2. Die Funktionsverläufe der Nutzungsgrade sind in Kapitel 5.2.2.1.3.1 beschrieben.

⁷⁶³ Vgl. Kapitel 5.2.2.1.3.1

⁷⁶⁴ Vgl. Kapitel 5.1.3.1; In Anlehnung an die Konzepte zur automatisierten Klassifizierung von Verma et al. und Mesnier et al. Der Lebenszyklus kann bspw. durch die Berechnung des arithmetischen Mittels der Nutzungsgrade der einzelnen Dateien errechnet werden. So entsteht ein Lebenszyklus, der den Nutzungsgrad der Daten eines Dateiordners widerspiegelt. Dies vereinfacht den Aufwand der Klassifizierung (Vgl. Bhagwan et al. / Management of Data Storage / ; Shah et al. / Classification for ILM / 12-14). Weiterhin ist denkbar, dass vor der Klassifizierung die Daten sogenannten Dokumentenklassen zugeordnet werden (Vgl. Kaiser, Smolnik, Riempp / ILM-Framework / 488). Die Grundlage hierfür bilden Geschäftsprozesse, in denen die Daten verwendet werden. Durch eine Zuordnung der Daten zu Dokumentenklassen oder Dateiordnern werden diese als zusammengehörig betrachtet und klassifiziert. Somit kann vermieden werden, dass zusammengehörige Daten durch Konzepte zur automatisierten Klassifizierung auf unterschiedlichen Speicherebenen abgelegt werden. Problematisch wäre bspw., wenn Teile der Daten eines Geschäftsprozesses in der Online-Ebene verwaltet werden, und wiederum andere Teile archiviert oder gelöscht werden. Dies könnte unter Umständen die ordnungsgemäße Ausführung des Geschäftsprozesses verhindern. Vgl. hierzu Kaiser, Smolnik, Riempp / ILM-Framework / 487-489. Weiterhin kann vermieden werden, dass bspw. Daten automatisch der Cloud-Ebene zugeordnet werden, obwohl dies aufgrund rechtlicher Anforderungen bedenklich wäre. Vgl. hierzu Kaplan, Rezek, Sprague / Information Protection Cloud / 16-17.

entstehen, werden die Funktionen der betrachteten sechs Lebenszyklen zusätzlich variiert.⁷⁶⁵ Durch den Einsatz von stochastischem Rauschen werden die Lebenszyklen mit Hilfe unterschiedlich starker Rauschgrade verändert.⁷⁶⁶ Wie die Daten stochastisch verauscht werden, wird im Folgenden erörtert. Zunächst werden für jeden Nutzungsgrad zwei im Intervall $[0,1]$ gleichverteilte unabhängige Zufallszahlen Z_1 und Z_2 erzeugt.⁷⁶⁷ Durch Transformation entsteht eine standard-normalverteilte Zufallszahl:⁷⁶⁸

$$N(0,1) = \sqrt{-2 \ln Z_1} \cdot \sin(2 \pi Z_2)$$

Der entsprechende Rauschgrad wird anschließend erzeugt durch:

$$N(\delta) = \delta \cdot N(0,1)$$

Durch setzen des δ -Parameters bzw. Rauschfaktors wird der Rauschgrad erhöht.⁷⁶⁹ Der resultierende Rauschgrad $N(\delta)$ wird anschließend zu dem jeweiligen Wert des Nutzungsgrads hinzuaddiert.⁷⁷⁰ Im Folgenden werden die generierten und verwendeten Le-

⁷⁶⁵ Damit werden nicht nur idealtypische Nutzungsgrade betrachtet. Darüber hinaus werden stochastische bzw. zufallsbasierte Veränderungen in den Nutzungsgraden in die Klassifizierung einbezogen.

⁷⁶⁶ Vgl. Box, Muller / Random Normal Deviates / 610-611; Nissen, Propach / Optimization Noise / 107-119; Nissen, Matthesius / Forecast Model Selection / 1-3

⁷⁶⁷ Die Erzeugung der Zufallszahlen mit unterschiedlichem Seed bzw. Startwert erfolgte in Microsoft Excel 2013 mit Hilfe der Funktion „Zufallszahl“ (Vgl. Charnes / Financial Modeling / 110-113). Die verschiedenen Lebenszyklen wurden je Simulationslauf mit unterschiedlichen Zufallszahlen und Seeds variiert. Bezogen auf die zwei betrachteten Szenarien (Vgl. Kapitel 5.2.2.1, Kapitel 5.2.3.1.1.1) werden für beide Szenarien die gleichen Lebenszyklen ohne und mit Rauschfaktoren verwendet. Dies ermöglicht den direkten Vergleich der Klassifizierungsergebnisse beider Szenarien. Vgl. hierzu Kapitel 5.3.

⁷⁶⁸ Vgl. Box, Muller / Random Normal Deviates / 610-611; Das arithmetische Mittel der standard-normalverteilten Zufallszahlen liegt nah beim Wert 0 und die Standardabweichung beträgt 1.

⁷⁶⁹ Die in der prototypischen Implementierung verwendeten Faktoren zur Veränderung der Rauschgrade betragen 5000, 10000 und 20000 für die jeweiligen Lebenszyklen (Vgl. Kapitel 5.2.2.1.3.1.1 bis 5.2.2.1.3.1.6). Diese Rauschgrade eignen sich zur Variation der Lebenszyklen. Für die Anzahl der Zugriffe pro Monat, die in den Lebenszyklen ohne Rauschfaktor abgebildet sind, wird ein Intervall von minimal 0 bis maximal 14400 Zugriffen festgelegt. Dieses Intervall wurde in Anlehnung an vorhandene Studien zu Lebenszyklen von Informationen definiert (Vgl. Gibson, Miller, Long / File-Usage / 641-643). Grundsätzlich können hierfür entsprechend der Zugriffe auf die Daten beliebige Intervalle festgelegt werden. Hierbei müssen lediglich die Rauschfaktoren angepasst werden, um angemessene Rauschgrade zu erzeugen.

⁷⁷⁰ Zur Sicherstellung statistischer Signifikanz werden für jeden Lebenszyklus und für jeden Rauschgrad 30 Simulationsläufe durchgeführt (Vgl. Sachs, Hedderich / Angewandte Statistik / 308-309; Nissen, Bietahn / Stochastische Optimierung / 117). Die Zufallszahlen Z werden dabei mit einem neuen Seed bzw. Startwert erzeugt. Für jeden Lebenszyklus wird ein Zeitraum von 5 Jahren bzw. 60 Monaten betrachtet. Es werden die 60 Einzelwerte der Nutzungsgrade für jeden Lebenszyklus durch stochastisches Rauschen verändert.

benszyklen graphisch dargestellt. Zu Darstellungszwecken der sechs Lebenszyklen in den Abbildungen 5-9 bis 5-14 wird von den einzelnen Rauschgraden sowie den daraus resultierenden Nutzungsgraden das arithmetische Mittel aus den einzelnen verrauschten bzw. variierten Lebenszyklen gebildet.⁷⁷¹ Somit wird verdeutlicht, wie stark sich ein bestimmter Rauschgrad auf den jeweiligen Lebenszyklus der Daten auswirkt. Für die später durchgeführten Simulationsexperimente wird nicht das arithmetische Mittel verwendet, sondern die jeweiligen Rauschgrade von 30 Simulationsläufen je Rauschgrad und Lebenszyklus. In den Abbildungen 5-9 bis 5-14 werden die sinkenden, steigenden, konstanten, saisonalen und trend-saisonalen Nutzungsgrade⁷⁷² vorgestellt.

5.2.2.1.3.1.1 Sinkender Nutzungsgrad

Der sinkende Nutzungsgrad ist in Abbildung 5-9 dargestellt.

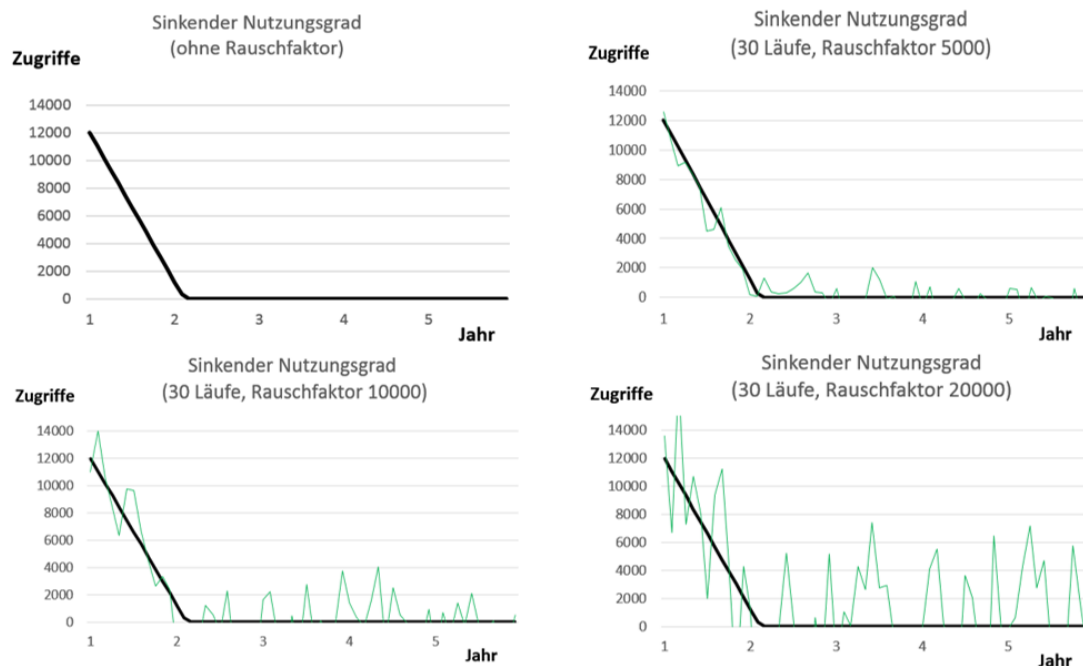


Abb. 5-9: Sinkender Nutzungsgrad ohne und mit Rauschfaktor

⁷⁷¹ Vgl. Kapitel 5.2.2.1.3.1.1 bis 5.2.2.1.3.1.6

⁷⁷² Die Datenreihen der Lebenszyklen ohne Rauschfaktor, die auch die Grundlage für die Lebenszyklen mit Rauschfaktor bilden, sind in Anhang 2 dargestellt. Wenn durch die Rauschfaktoren negative Nutzungsgrade entstehen, so werden diese auf den Wert 0 gesetzt.

Die häufigsten Zugriffe erfolgen erfahrungsgemäß kurz nach dem Entstehen der Information.⁷⁷³ Danach wird in der Regel nur noch selten auf die Information zugegriffen. Dies ist bspw. bei E-Mails der Fall.⁷⁷⁴ Gleiches gilt für Marketingdaten, wie bspw. Produktwerbungen oder Preisaktionen. Diese verlieren aufgrund der immer kürzer werdenden Produktlebenszyklen⁷⁷⁵ schnell an Bedeutung. Der Nutzungsgrad fällt im ersten Jahr rapide ab. Kurz nach dem zweiten Jahr wird nicht mehr auf die Daten zugegriffen.⁷⁷⁶ Der sinkende Nutzungsgrad wird durch den Einsatz von stochastischem Rauschen mit Hilfe von verschiedenen Rauschfaktoren verändert.⁷⁷⁷ Dies ist ebenfalls in Abbildung 5-9 dargestellt. Die in schwarzer Farbe abgebildete Funktion bildet jeweils den Lebenszyklus ohne Rauschgrad ab und die grüne Funktion die jeweils verrauschten Lebenszyklen.⁷⁷⁸ Die Funktion ohne Rauschgrad bildet die Grundlage für die Funktionen mit Rauschfaktor. In der Realität existieren keine stetigen Nutzungsgrade von Informationen, da die Zugriffe zufällig erfolgen. Dieser Sachverhalt wird durch die Funktionen mit Rauschgrad abgebildet. Es wird deutlich, dass bei stärkerem Rauschfaktor die Zugriffe im Zeitverlauf bzw. der Nutzungsgrad stärker variieren.

⁷⁷³ Vgl. Chen / Information valuation / 135-146; Zadok et al. / Reducing Storage Management Costs / 101-105)

⁷⁷⁴ Vgl. Bennett, Bauer, Kinchlea / NFS environments / 18-25; Chen / Information valuation / 135-146; Inmon / ILM Data Warehousing / 1-20; Shah et al. / Classification for ILM / 1-15; Short / ILM / 2-37; Zadok et al. / Reducing Storage Management Costs / 101-105.

⁷⁷⁵ Vgl. Burmann, Freiling, Hülsmann / Ad-hoc-Krisen / 5; Kuhn, Hellingrath / SCM / 33

⁷⁷⁶ Vgl. Kapitel 2.4.2 zur Beschreibung von Beispielen von unstrukturierten Daten mit sinkendem Nutzungsgrad.

⁷⁷⁷ Vgl. Kapitel 5.2.2.1.3.1

⁷⁷⁸ Vgl. Kapitel 5.2.2.1.3.1; Abbildung 5-9 bis Abbildung 5-14

5.2.2.1.3.1.2 Steigender Nutzungsgrad

Abbildung 5-10 zeigt einen steigenden Nutzungsgrad der Informationen. Nach dem Entstehen der Daten nimmt der Nutzungsgrad ab. Ab dem zweiten Jahr nehmen die Zugriffe auf die Informationen zu.⁷⁷⁹ Dies ist bspw. bei Daten der Forschung und Entwicklung denkbar, wenn Forschungsergebnisse im Laufe der Zeit wichtiger werden und in Produkte umgesetzt werden können.⁷⁸⁰ Ein weiteres Beispiel hierfür ist ein Internet-Portal, das seinen Nutzern den Austausch von digitalen Inhalten, wie Bildern oder Videos, ermöglicht. Liegen bestimmte Inhalte besonders im Trend, greifen die Nutzer verstärkt auf die Inhalte zu. Vereinzelt ist dabei auch mit sehr hohen Nutzungsgraden zu bestimmten Zeitpunkten zu rechnen, mit sogenannten Lastspitzen.⁷⁸¹

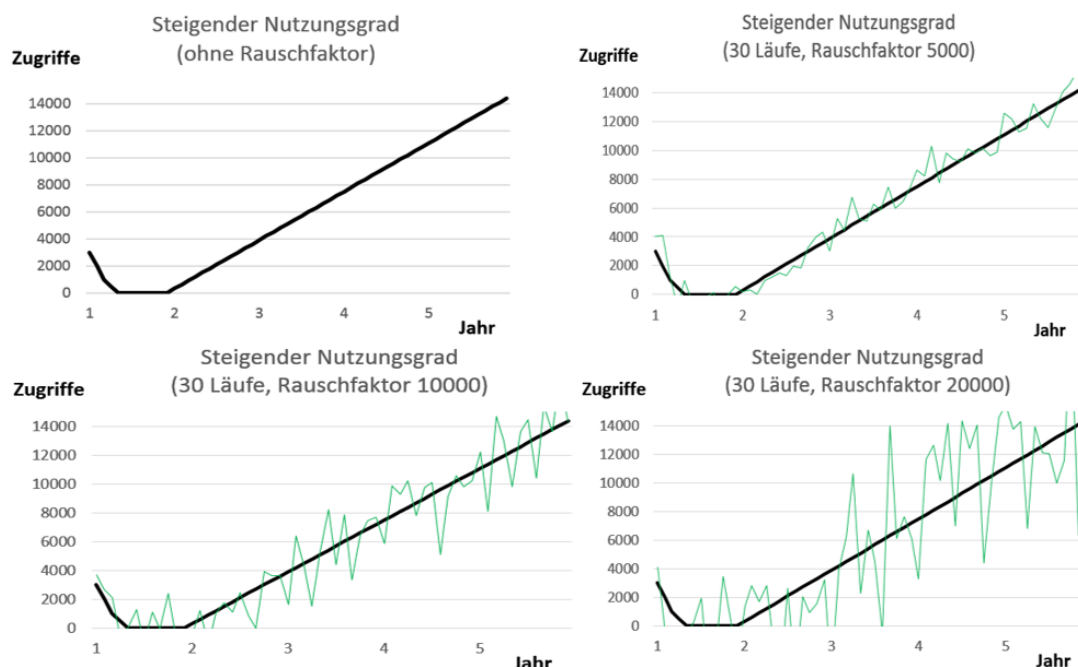


Abb. 5-10: Steigender Nutzungsgrad ohne und mit Rauschfaktor

⁷⁷⁹ Vgl. Kapitel 2.4.2 zur Beschreibung von Beispielen von unstrukturierten Daten mit steigendem Nutzungsgrad.

⁷⁸⁰ Vgl. Lundberg / Lifecycle Management / 5

⁷⁸¹ Vgl. Shelly, Rosenblatt / Systems Design / 12; Armbrust et al. / Cloud Computing / 9-11; Baun et al. / Cloud Computing / 63-64; Kapitel 3.2.2

Der steigende Nutzungsgrad aus Abbildung 5-10 wird durch den Einsatz von stochastischem Rauschen mit Hilfe von verschiedenen Rauschfaktoren verändert.⁷⁸² Die veränderten Nutzungsgrade sind in grüner Farbe in Abbildung 5-10 dargestellt.

5.2.2.1.3.1.3 Konstanter Nutzungsgrad

Abbildung 5-11 zeigt einen konstanten Nutzungsgrad der Informationen. Die Zugriffe bleiben über den betrachteten Zeitraum von fünf Jahren konstant. Konstante Zugriffe erfolgen bspw. bei Logdateien oder systemseitig gespeicherten Daten.⁷⁸³ Diese Daten werden benötigt, um bspw. die Funktionalität von Betriebssystemen nutzen zu können.⁷⁸⁴

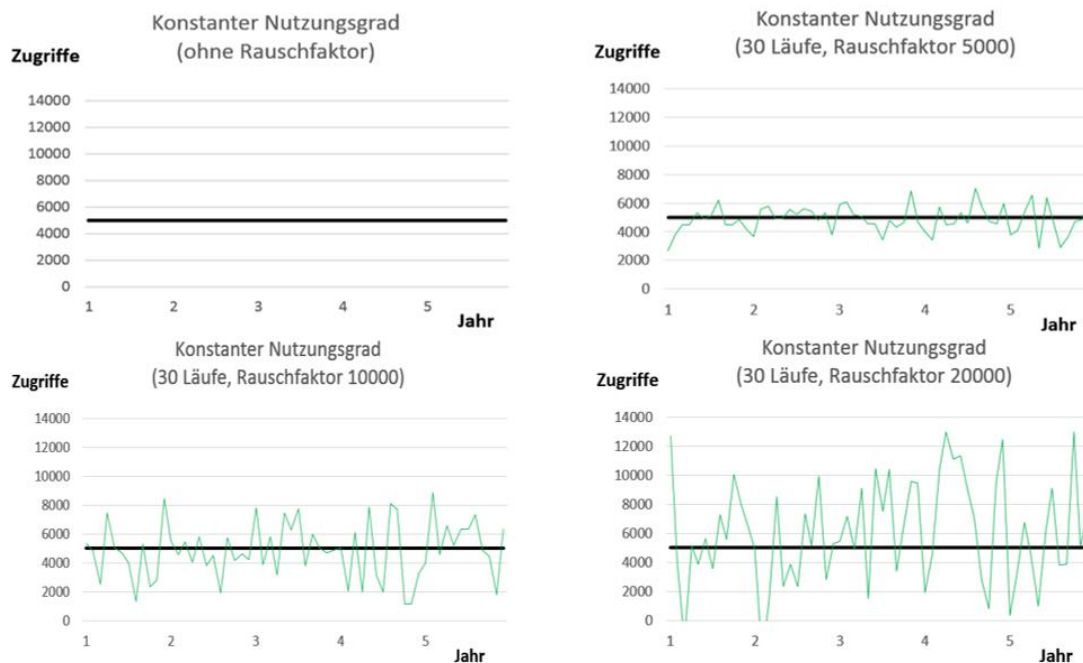


Abb. 5-11: Konstanter Nutzungsgrad ohne und mit Rauschfaktor

⁷⁸² Vgl. Kapitel 5.2.2.1.3.1

⁷⁸³ Vgl. Fairbanks / Software Architecture / 4

⁷⁸⁴ Vgl. Verma et al. / Very Large File Systems / 160-168

5.2.2.1.3.1.4 Saisonaler Nutzungsgrad

Abbildung 5-12 zeigt einen saisonalen bzw. periodisch variierenden Nutzungsgrad der Informationen. Saisonal bedingte Zugriffe erfolgen bspw. auf Finanzdaten. Aufgrund der Erstellung von Quartalsabschlüssen oder Jahresberichten erfolgt in der Regel ein periodischer Zugriff auf Finanzdaten.⁷⁸⁵ Zeitweise erfolgen keine Zugriffe auf die Informationen.

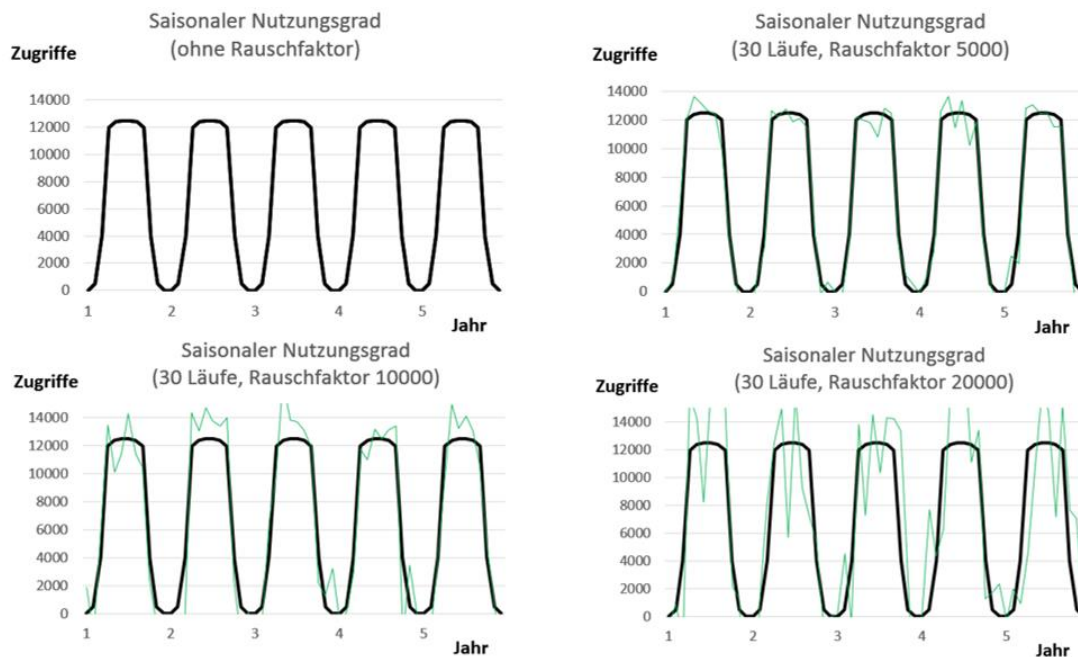


Abb. 5-12: Saisonaler Nutzungsgrad ohne und mit Rauschfaktor

5.2.2.1.3.1.5 Sinkender Trend-Saisonaler Nutzungsgrad

Abbildung 5-13 zeigt einen saisonal variierenden Nutzungsgrad der Informationen mit einer sinkenden Trendkomponente. In Internet-Plattformen erfolgen bspw. jahreszeitlich bedingte saisonale Zugriffe auf die Daten. Sinken die Zugriffe auf die Daten im Zeitverlauf, liegen sinkende trend-saisonale Nutzungsgrade vor.⁷⁸⁶ Beispiele hierfür sind Internet-Plattformen, die Produkte verkaufen. Die Produkte werden saisonal bedingt von den Käufern nachgefragt. Es erfolgen Zugriffe auf die Produktdaten, um relevante produktbezogene Informationen einzuholen oder das Produkt zu kaufen. Mit Voranschreiten

⁷⁸⁵ Vgl. Kapitel 2.4.2

⁷⁸⁶ Vgl. Armbrust et al. / Cloud Computing / 10

des Produktlebenszyklus einer Produktlinie wird die Nachfrage nach diesen Produkten im Zeitverlauf geringer, der Zugriff auf die Produktdaten sinkt.⁷⁸⁷

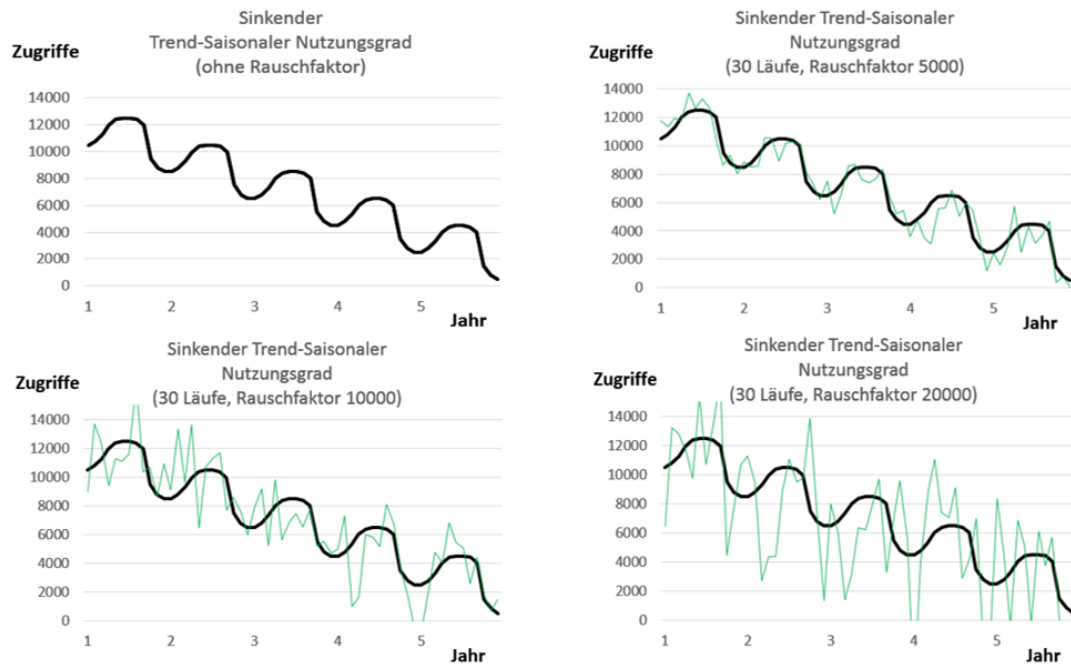


Abb. 5-13: Sinkender Trend-Saisonaler Nutzungsgrad ohne und mit Rauschfaktor

Der saisonale Nutzungsgrad wird ebenfalls durch verschiedene Rauschfaktoren verändert.⁷⁸⁸ Je stärker die Rauschfaktoren sind, desto unregelmäßiger sind die Zugriffe auf die Daten im Zeitverlauf. Insbesondere bei Internet-basierten Plattformen sind die Zugriffe auf die Daten schwer vorhersehbar und sehr unregelmäßig. Diese unterliegen starken Schwankungen.⁷⁸⁹

⁷⁸⁷ Vgl. Shelly, Rosenblatt / Systems Design / 12

⁷⁸⁸ Vgl. Kapitel 5.2.2.1.3.1

⁷⁸⁹ Vgl. hierzu Armbrust et al. / Cloud Computing / 10 und Kapitel 3.2.2, Kapitel 3.4.1.

5.2.2.1.3.1.6 Steigender Trend-Saisonaler Nutzungsgrad

Abbildung 5-14 zeigt einen saisonal variierenden Nutzungsgrad der Informationen mit einer steigenden Trendkomponente. Denkbar wären saisonal oder jahreszeitlich bedingte Zugriffe auf die Daten einer Internet-Plattform. Steigen die Zugriffe auf die Daten im Zeitverlauf, liegen steigende trend-saisonale Nutzungsgrade vor.

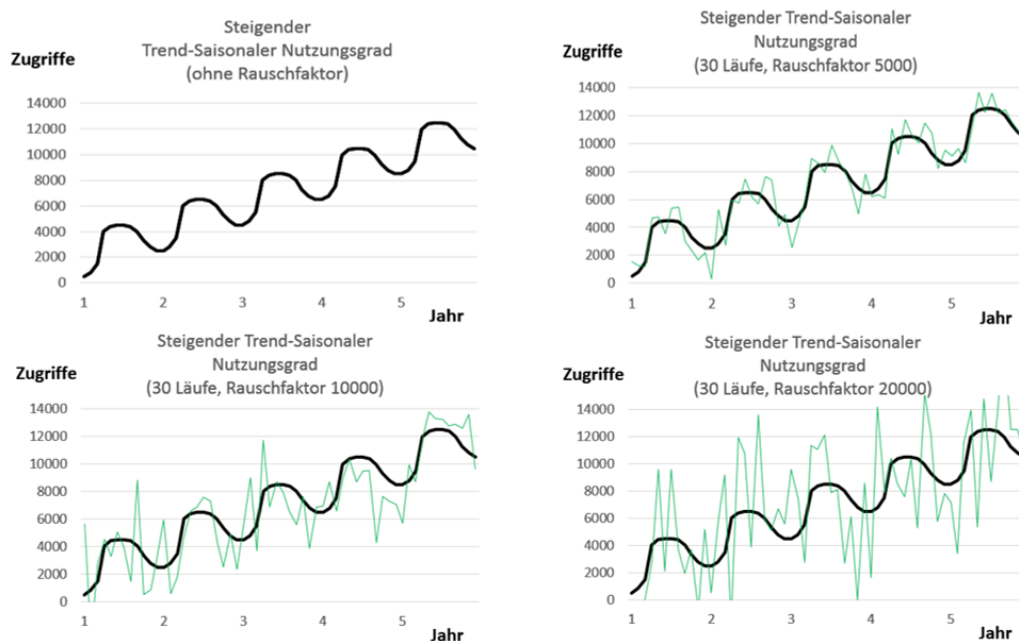


Abb. 5-14: Steigender Trend-Saisonaler Nutzungsgrad ohne und mit Rauschfaktor

Ein Beispiel hierfür sind saisonal bedingte Zugriffe auf Daten von Produkten, die über eine Internet-Plattform vertrieben werden und die im Laufe der Zeit, bspw. durch das Entstehen von Modetrends, stärker nachgefragt werden.⁷⁹⁰

⁷⁹⁰ Vgl. Armbrust et al. / Cloud Computing / 8-11

5.2.3 Implementierung eines Verfahrens zur Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM

5.2.3.1 Klassifizierungsfunktionen

5.2.3.1.1 Implementierung der Klassifizierungsfunktionen

Durch die Implementierung und Anwendung von Klassifizierungsfunktionen⁷⁹¹ werden die Informationen entsprechend ihres Nutzungsgrads⁷⁹² einer bestimmten Klasse zugeordnet.⁷⁹³ Der Aufbau und die Funktionsweise der Klassifizierungsfunktionen werden in diesem Kapitel vorgestellt. Es werden zwei Szenarien zur Klassifizierung der Daten betrachtet.⁷⁹⁴ Im ersten Szenario werden ausschließlich unternehmensinterne Speichermedien bzw. Speicherebenen berücksichtigt. Im zweiten Szenario wird neben den unternehmensinternen Speichermedien zusätzlich die Cloud-Ebene⁷⁹⁵ in die Betrachtungen einbezogen.⁷⁹⁶ Für jedes Szenario wird eine Klassifizierungsfunktion in Microsoft Excel implementiert, die entsprechend der Nutzungsgrade die Informationen einer Klasse zuordnet.⁷⁹⁷ Hierzu wird die WENN-Funktion⁷⁹⁸ verwendet. Die Syntax der WENN-Funktion lautet wie folgt:

= WENN (Prüfung; [Dann_Wert]; [Sonst_Wert])

Die WENN-Funktion kann im dargestellten Beispiel zwei Werte zurückliefern. Der erste Wert wird zurückgegeben, wenn eine festgelegte Bedingung „wahr“ ist.⁷⁹⁹ Der zweite Wert wird zurückgegeben, wenn die Bedingung „falsch“ ist.⁸⁰⁰ Zur Implementie-

⁷⁹¹ Vgl. Kapitel 5.2.2.1.2

⁷⁹² Vgl. Kapitel 5.2.2.1.3.1

⁷⁹³ Vgl. Kapitel 5.1.2.3

⁷⁹⁴ Vgl. Kapitel 5.2.2.1, Kapitel 5.2.3.1.1.1.1, Kapitel 5.2.3.1.1.1.2

⁷⁹⁵ Vgl. Kapitel 5.1.2.3

⁷⁹⁶ Vgl. Kapitel 5.2.3.1.1

⁷⁹⁷ Vgl. Anhang 4 zur Implementierung der Klassifizierungsfunktion in Microsoft Excel.

⁷⁹⁸ Vgl. Kofler, Nebelo / Excel Programmieren / 480-481. Mit Hilfe der WENN-Funktion wird die Kausalbeziehung zwischen Input- und Output-Größen bzw. „wenn-dann“-Beziehungen abgebildet. Diese Vorgehensweise eignet sich für die Simulation komplexer Vorgänge, beispielsweise für die Abbildung stochastischer Simulationsmodelle. Vgl. hierzu Küll, Stähly / Simulationsexperimente / 1-4.

⁷⁹⁹ Hierbei handelt es sich um den „Dann_Wert“.

⁸⁰⁰ Hierbei handelt es sich um den „Sonst_Wert“.

nung der Klassifizierungsfunktionen mit Hilfe der WENN-Funktion werden die Parameter (n), (i) und (k) verwendet. Der Nutzungsgrad bzw. der Zugriff auf eine Datei zu einem bestimmten Zeitpunkt, wird durch (n) beschrieben. Das Intervall (i) gibt an, ab welchem Minimalwert und bis zu welchem Maximalwert, bezüglich des Nutzungsgrads, die Daten einer bestimmten Klasse angehören sollen. Die Existenz mehrerer Intervalle (i_1) bis (i_x) innerhalb einer Klassifizierungsfunktion ist zulässig.⁸⁰¹ Der Parameter (k) bezeichnet die zum Intervall (i) zugehörige Klasse.⁸⁰² Durch Übertragung dieser Parameter auf die WENN-Funktion entsteht die Klassifizierungsfunktion K (n, i, k).⁸⁰³ Die Syntax der Klassifizierungsfunktion K (n, i, k) lautet wie folgt:

$$K(n, i, k) = (\text{WENN}(n_t \geq i_{1[\min]}; [k_1]; [k_2]))$$

Die obige Funktion prüft, ob der Nutzungsgrad (n) zum Zeitpunkt t dem Intervall (i_1) angehört.⁸⁰⁴ Ist die Bedingung erfüllt, gehört (n) dem Intervall (i_1) an.⁸⁰⁵ Somit werden die Daten mit dem Nutzungsgrad (n) der Klasse (k_1) zugeordnet. Wenn (n) dem Intervall (i_1) nicht angehört, werden die Daten der Klasse (k_2) zugeordnet.⁸⁰⁶ Dieses Beispiel ist anwendbar, wenn zwei Intervalle (i_1) und (i_2) sowie zwei zugehörige Klassen (k_1) und (k_2) existieren.⁸⁰⁷ Da in dieser Arbeit mehr als zwei Intervalle und Klassen existie-

⁸⁰¹ Bei Aufteilung in mehrere Intervalle kann der Maximalwert des zu prüfenden Nutzungsgrads für das erste Intervall i_1 beliebig hoch sein. Der Minimalwert x des zuletzt vorkommenden Intervalls i_x beträgt 0. Dieser Wert repräsentiert den Nutzungsgrad 0, wenn keine Zugriffe auf die Daten erfolgen.

⁸⁰² Der Parameter (k) liefert den Wert der entsprechenden Klasse zurück, bspw. für das erste Szenario den Wert „3“ für die Klasse „Online“, den Wert „2“ für die Klasse „Nearline“ und den Wert „1“ für die Klasse „Offline“.

⁸⁰³ Die Klassifizierungsfunktion K wird im Formeleditor von Microsoft Excel erzeugt. Vgl. hierzu Anhang 4.

⁸⁰⁴ Der Nutzungsgrad (n) kann im Minimum den Wert 0 annehmen, wenn keine Zugriffe auf die Daten erfolgen. Für den maximal annehmbaren Wert des Nutzungsgrads existiert keine definierte Grenze im Rahmen der Klassifizierungsfunktion. Dieser Wert wird lediglich limitiert durch den höchsten Nutzungsgrad innerhalb eines Lebenszyklus. Da die Klassifizierung basierend auf den Monatswerten der Nutzungsgrade erfolgt (Vgl. Kapitel 5.2.2.1.2), handelt es sich beim Zeitpunkt (t) um den jeweils betrachteten Monat.

⁸⁰⁵ Ist der Wert des Nutzungsgrads größer als der Minimalwert des Intervalls oder entspricht er dem Minimalwert des Intervalls, so wird der Wert des Nutzungsgrads dem Intervall (i_1) zugeordnet.

⁸⁰⁶ Die Klassifizierungsfunktion muss nicht gesondert prüfen, ob der Nutzungsgrad (n) dem Intervall (i_2) angehört. Wenn der Nutzungsgrad dem ersten Intervall nicht zugehörig ist, ist die Zugehörigkeit zu (i_2) automatisch erfüllt, da die größer oder gleich Bedingung bezüglich des Minimalwerts von Intervall (i_1) nicht erfüllt ist und der Wertebereich des Intervalls (i_2) kleiner ist als der Wertebereich des Intervalls (i_1).

⁸⁰⁷ Denkbar wäre bspw. die Existenz der Klassen „2“ und „1“ bzw. „Online“ und „Nearline“ und die jeweils dazugehörigen Intervalle zur Bestimmung der Anzahl der Nutzungsgrade. Vgl. Anhang 4 zur Implementierung der Klassifizierungsfunktion in Microsoft Excel. Hinsichtlich der Wertebereiche bzw.

ren, denen die Daten zugeordnet werden, muss die Klassifizierungsfunktion erweitert werden.⁸⁰⁸ Hierzu werden die Klassifizierungsfunktionen verschachtelt. Im ersten Szenario⁸⁰⁹ existieren drei unterschiedliche Klassen (k). Die Klassifizierungsfunktion $K_1(n, i, k)$ ergibt sich wie folgt:

$$K_1(n, i, k) = (\text{WENN}(n_t \geq i_{1 \text{ [min]}}; [k_1]; (\text{WENN}(n_t \geq i_{2 \text{ [min]}}; [k_2]; [k_3])))$$

Die Funktion $K_1(n, i, k)$ prüft, ob der Nutzungsgrad (n) innerhalb der zwei Intervalle (i_1) und (i_2) enthalten ist.⁸¹⁰ Wenn nicht, dann werden die Daten automatisch der dritten Klasse (k_3) zugeordnet.⁸¹¹ Im zweiten Szenario⁸¹² existieren insgesamt vier Intervalle und Klassen, sodass die Klassifizierungsfunktion für dieses Szenario nochmals erweitert werden muss. Die Klassifizierungsfunktion $K_2(n, i, k)$ lautet wie folgt:

$$K_2(n, i, k) = (\text{WENN}(n_t \geq i_{1 \text{ [min]}}; [k_1]; (\text{WENN}(n_t \geq i_{2 \text{ [min]}}; [k_2]; (\text{WENN}(n_t \geq i_{3 \text{ [min]}}; [k_3]; [k_4]))))))$$

Durch die Erweiterung der Klassifizierungsfunktion wird mit der Funktion $K_2(n, i, k)$ die Klassifizierung und anschließende Zuordnung der Daten in vier unterschiedliche Klassen ermöglicht.⁸¹³ In Anhang 3 ist die Funktionsweise der Klassifizierungsfunktionen K_1 und K_2 jeweils als Flussdiagramm dargestellt.

Größe der einzelnen Intervalle gilt: $(i_1) > (i_2)$. Vgl. hierzu Kapitel 5.2.3.1.1.1.

⁸⁰⁸ Bezogen auf die zwei betrachteten Szenarien zur Klassifizierung unstrukturierter Daten (Vgl. hierzu Kapitel 5.2.3.3.3). Die Daten werden im ersten Szenario drei Klassen und im zweiten Szenario vier Klassen zugeordnet.

⁸⁰⁹ Vgl. Kapitel 5.2.3.1 und Kapitel 5.3

⁸¹⁰ Dies wird durch die in der Klassifizierungsfunktion enthaltenen beiden Terme „ $n_t \geq i_{1 \text{ [min]}}$ “ und „ $n_t \geq i_{2 \text{ [min]}}$ “ überprüft.

⁸¹¹ Hinsichtlich der Wertebereiche bzw. Größe der einzelnen Intervalle gilt: $(i_1) > (i_2) > (i_3)$. Vgl. hierzu Kapitel 5.2.3.1.1.1.

⁸¹² Vgl. Kapitel 5.2.3.1; Kapitel 5.2.3.1.1.1.2

⁸¹³ Hinsichtlich der Wertebereiche bzw. Größe der vier Intervalle gilt: $(i_1) > (i_2) > (i_3) > (i_4)$. Vgl. hierzu Kapitel 5.2.3.1.1.1.

5.2.3.1.1.1 Szenarien zur Klassifizierung unstrukturierter Daten

Es werden zwei Szenarien zur Klassifizierung unstrukturierter Daten untersucht. Im ersten Szenario wird die Klasse bzw. Speicherebene „Cloud“⁸¹⁴ nicht betrachtet. Aufgrund der in den Kapiteln 2.4.3 und 3.4.2 erörterten rechtlichen Anforderungen kann die Nutzung von unternehmensexternen Speicherebenen bzw. Public-Cloud-Diensten⁸¹⁵ ein Risiko hinsichtlich der Vertraulichkeit der Informationen darstellen. Es wird angenommen, dass die Informationen im ersten Szenario ausschließlich auf unternehmensinternen Speichermedien verwaltet werden sollen. Im zweiten Szenario werden ebenfalls unternehmensinterne Speichermedien zur Verwaltung der Informationen verwendet. Darüber hinaus werden für Informationen, die einen hohen Nutzungsgrad aufweisen, Public-Cloud-Dienste bzw. IaaS-Dienste eingesetzt.⁸¹⁶ Durch den Vergleich der Klassifizierungsergebnisse beider Szenarien kann untersucht werden, inwiefern die zusätzliche Nutzung von Cloud-Diensten⁸¹⁷ Auswirkungen auf das Kosten- und Nutzenverhältnis bezüglich der Verwaltung unstrukturierter Daten im ILM hat.⁸¹⁸

5.2.3.1.1.1.1 Szenario 1: Klassifizierung der Daten ohne Cloud-Ebene

Die Funktion $K_1(n, i, k)$ prüft, ob der Nutzungsgrad (n) innerhalb der zwei Intervalle (i_1) und (i_2) enthalten ist.⁸¹⁹ Befindet sich der Nutzungsgrad in einem der beiden Intervalle, wird den Klassen der Wert „3“ oder „2“ zugeordnet. Ist dies nicht der Fall, erfolgt die Zuordnung des Werts „1“.⁸²⁰ Die Klassifizierungsfunktion ergibt sich wie folgt:

⁸¹⁴ Vgl. Kapitel 3.4.2

⁸¹⁵ Vgl. Kapitel 3.2.1

⁸¹⁶ Dies ist bspw. sinnvoll zur Abdeckung von Lastspitzen bzw. stark variierenden Nutzungsgraden, da die Speicherressourcen in Abhängigkeit von der Nutzung mit Hilfe einer Public-Cloud skaliert werden können (Vgl. Kapitel 3.2.1 und 3.2.2; Armbrust et al. / Cloud Computing / 10-12). Sinken die Nutzungsgrade, können die Informationen mit Hilfe der unternehmensinternen Speichermedien verwaltet werden oder entsprechend rechtlicher Anforderungen langfristig aufbewahrt werden.

⁸¹⁷ In dieser Arbeit wird angenommen, dass zur Verwaltung der unstrukturierten Daten in der Cloud der Simple Storage Service (Amazon S3) vom Anbieter Amazon verwendet wird. Amazon S3 ist der derzeit bekannteste und am weitesten verbreitete Dienst für Cloud-Speicher. Vgl. hierzu Magoulès, Pan, Teng / Data Intensive Computing / 12; Darcey, Conder / Android Application Programming / 227; Buyya, Vecchiola, Selvi / Cloud Computing / 92; Cusumano / SaaS and Cloud Computing / 10.

⁸¹⁸ Vgl. Kapitel 5.3

⁸¹⁹ Vgl. Kapitel 5.2.3.1.1

⁸²⁰ Die Klassen mit den Werten „3“, „2“ und „1“ repräsentieren jeweils die Klassen „Online“, „Nearline“ und „Offline“.

$$\mathbf{K_1(n, i, k)} = (\text{WENN}(n_t \geq i_{1[\min]}; [k_1]; (\text{WENN}(n_t \geq i_{2[\min]}; [k_2]; [k_3])))$$

für $n \in [0; x]$,

$i_{1[\min, \max]} \in [3000; x]$ und $k_1 = [3]$,

$i_{2[\min, \max]} \in [1; 2999]$ und $k_2 = [2]$,

$i_{3[\min, \max]} \in [0; 0]$ und $k_3 = [1]$.

Die Klassifizierung der im Kapitel 5.2.2.1.3.1 vorgestellten Lebenszyklen bzw. Nutzungsgrade erfolgt mit Hilfe von Microsoft Excel 2013.⁸²¹ Die Größe bzw. die Wertebereiche der jeweiligen Klassenintervalle wurden im Rahmen der prototypischen Implementierung anhand der Lese- und Schreibgeschwindigkeiten der zugehörigen Speichermedien⁸²² festgelegt.⁸²³ Informationen, die einen hohen Nutzungsgrad aufweisen, sollen auf Speichern mit kurzen Zugriffszeiten⁸²⁴ verwaltet werden, die dementsprechend kostenintensiv sind. Deshalb werden diese Informationen der Klasse „Online“ zugeordnet. Informationen, die einen geringen Nutzungsgrad aufweisen, werden der Klasse „Nearline“ zugeordnet. Erfolgen keine Zugriffe auf die Informationen, erfolgt die Zuordnung zur Klasse „Offline“. Zwischen den Klassen bzw. Speicherebenen bestehen signifikante Unterschiede bezüglich der Zugriffszeiten und Übertragungsraten.⁸²⁵

⁸²¹ Die Werte für den Nutzungsgrad, die Intervalle und zugehörigen Klassennamen verweisen dabei auf die Zellen in Microsoft Excel, welche die zugehörigen Werte enthalten. In Anhang 4 wird beispielhaft dargestellt, wie die Intervalle, Klassen und Klassifizierungsfunktionen implementiert wurden. Wie in Kapitel 5.2.2.1.2 beschrieben, wird die Klassifizierung für jeden Monatswert der Nutzungsgrade durchgeführt. Dabei werden die Nutzungsgrade der Informationen des jeweiligen Vormonats in die Klassifizierung einbezogen. Jeder der betrachteten Lebenszyklen (Vgl. Kapitel 5.2.2.1.3.1) besteht aus 60 Monatswerten (Vgl. Anhang 1). Die Klassifizierungsfunktion wird also für jeden Lebenszyklus ohne Rauschfaktor je Szenario (Vgl. Kapitel 5.2.3.1.1.1.1, Kapitel 5.2.3.1.1.1.2) 60 Mal aufgerufen. Für die 6 betrachteten Lebenszyklen ohne Rauschfaktor ergeben sich je Szenario 360 Funktionsaufrufe. Für die Lebenszyklen mit Rauschfaktoren werden 30 Läufe je Lebenszyklus und Rauschfaktor simuliert (Vgl. Kapitel 5.2.2.1.3.1). Bezogen auf die 6 Lebenszyklen und 3 Rauschfaktoren ergeben sich im Rahmen der Simulation 32.400 Aufrufe der Klassifizierungsfunktion je Szenario.

⁸²² Vgl. Kapitel 2.4.2, Kapitel 5.1.2.4

⁸²³ Die Intervalle sind entsprechend der Nutzungsgrade und eingesetzten Speicherebenen festzulegen und demzufolge unternehmensspezifisch. Vgl. hierzu Dannehl / Speicherbedarf / 8. Im Rahmen dieser Arbeit werden geeignete Intervalle, basierend auf den Lebenszyklen und betrachteten Klassen, definiert.

⁸²⁴ Vgl. Kapitel 2.4.2

⁸²⁵ Vgl. hierzu Petrocelli / Information Lifecycle Management / 23; Maier, Hädrich, Peinl / Knowledge Infrastructures / 116

In der Online-Ebene werden Festplatten mit kurzen Zugriffszeiten und hohen Übertragungsraten verwendet.⁸²⁶ Aufgrund der kurzen Zugriffszeiten erfolgt eine Einordnung der Daten in die Klasse „Online“ ab einem Nutzungsgrad von 3000. Der Nutzungsgrad der Online-Ebene kann einen beliebig großen Wert annehmen, dargestellt durch das „x“ als obere Intervallgrenze.⁸²⁷ Die Klasse „Online“ hat somit den größten Wertebereich bzw. bildet das größte Teilintervall. Auf diese Informationen kann jederzeit schnell zugegriffen werden. Magnetplatten mit mittleren Zugriffszeiten werden in der Nearline-Ebene eingesetzt.⁸²⁸ Zeitkritische Anfragen an Informationen der Nearline-Ebene stellen eher die Ausnahme dar.⁸²⁹ Aufgrund der im Vergleich zu Festplatten deutlich längeren Zugriffszeiten erfolgt eine Einordnung der Daten in die Klasse „Nearline“ bis zu einem Nutzungsgrad von 2999. Die Klasse „Nearline“ bildet somit ein kleineres Teilintervall als die Klasse „Online“ ab. Für die Offline-Ebene werden kostengünstige Magnetbänder⁸³⁰ verwendet. Die Speicherung auf Magnetband gilt zwar als kostengünstigste Alternative,⁸³¹ allerdings ist der Zugriff auf diese Daten aufgrund der langen Zugriffszeiten⁸³² mit einem hohen Zeitaufwand verbunden. Aus diesem Grund erfolgt die Einordnung der Daten in die Klasse „Offline“ nur, wenn keine Zugriffe auf die Daten erfolgen, bzw. der Nutzungsgrad null ist.

5.2.3.1.1.2 Szenario 2: Klassifizierung der Daten mit Cloud-Ebene

Die Funktion $K_2(n, i, k)$ prüft im zweiten Szenario, ob der Nutzungsgrad (n) innerhalb der drei Intervalle (i_1), (i_2) oder (i_3) enthalten ist.⁸³³ Befindet sich der Nutzungsgrad

⁸²⁶ Mit Zugriffszeiten zwischen 2 und 4 Millisekunden und Übertragungsraten mit mehreren hundert Megabyte pro Sekunde. Vgl. hierzu Petrocelli / Information Lifecycle Management / 23; Maier, Hädrich, Peinl / Knowledge Infrastructures / 116.

⁸²⁷ Die maximalen Nutzungsgrade der betrachteten Lebenszyklen liegen in Bezug auf die Mittelwerte bei den Lebenszyklen mit dem größten Rauschfaktor bei ca. 21000.

⁸²⁸ Die Zugriffszeiten liegen in einem Bereich von 10-50 Sekunden (Vgl. hierzu Petrocelli / Information Lifecycle Management / 23; Rautenstrauch, Schulze / Informatik / 40).

⁸²⁹ Vgl. Gray, Reuter / Transaction Processing / 54

⁸³⁰ Vgl. Janko / Informationswirtschaft / 112-116

⁸³¹ Vgl. Hillyer, Silberschatz / Storage Systems / 195-204

⁸³² Vgl. Härder, Rahm / Datenbanksysteme / 67-68

⁸³³ Vgl. Kapitel 5.2.3.1.1; Gehört der Nutzungsgrad einer der drei Intervalle an, erfolgt die Zuordnung zu den Klassen „4“, „3“ oder „2“ bzw. „Cloud“, „Online“ oder „Nearline“.

nicht innerhalb dieser drei Intervalle, werden die Daten der vierten Klasse zugeordnet.⁸³⁴ Die Klassifizierungsfunktion lautet wie folgt:

$$\mathbf{K}_2(\mathbf{n}, \mathbf{i}, \mathbf{k}) = (\text{WENN} (n_t \geq i_{1 [\min]}; [k_1]; (\text{WENN} (n_t \geq i_{2 [\min]}; [k_2]; (\text{WENN} (n_t \geq i_{3 [\min]}; [k_3]; [k_4]))))))))$$

für $n \in [0; x]$,

$i_{1 [\min, \max]} \in [8000; x]$ und $k_1 = [4]$,

$i_{2 [\min, \max]} \in [3000; 7999]$ und $k_2 = [3]$,

$i_{3 [\min, \max]} \in [1; 2999]$ und $k_3 = [2]$,

$i_{4 [\min, \max]} \in [0; 0]$ und $k_4 = [1]$.

Die Klassifizierung der in Kapitel 5.2.2.1.3.1 vorgestellten Lebenszyklen bzw. Nutzungsgrade für das zweite Szenario erfolgt ebenfalls mit Hilfe von Microsoft Excel 2013. Die Größe der jeweiligen Klassenintervalle wurde im Rahmen der prototypischen Implementierung, wie in Kapitel 5.2.3.1.1.1.1 beschrieben, anhand der Lese- und Schreibgeschwindigkeiten der zugehörigen Speichermedien bzw. -ebenen festgelegt.⁸³⁵ Zusätzlich existiert für Szenario 2 die Klasse „Cloud“.⁸³⁶ Die Informationen, die einen hohen und stark variierenden Nutzungsgrad aufweisen, werden in diesem Szenario der Klasse „Cloud“ zugeordnet.⁸³⁷ Im Rahmen des zweiten Szenarios werden die Informationen ab einem Nutzungsgrad von 8000 der Klasse Cloud zugeordnet. Durch die Nutzung von Public-Cloud-Diensten bzw. Cloud-Speicher (IaaS) können Lastspitzen hinsichtlich des Nutzungsgrads abgedeckt werden und es werden kurze Zugriffszeiten im Millisekunden-Bereich ermöglicht.⁸³⁸ Die obere Grenze des Intervalls der Cloud-Ebene

⁸³⁴ Die vierte Klasse wird repräsentiert durch den Zahlenwert „1“ und steht für die „Offline“-Daten.

⁸³⁵ In Kapitel 5.3 wird zusätzlich erörtert, inwiefern sich die Modifikation der Intervalle der einzelnen Klassen auf das Kosten- und Nutzenverhältnis zur Verwaltung der Daten auswirkt.

⁸³⁶ Vgl. Kapitel 5.2.3.1.1.1; Vgl. zur Beschreibung der Eigenschaften der Speichermedien Kapitel 5.2.3.1.1.1.1.

⁸³⁷ Verschiedene Anbieter von Cloud-Speicher bieten zusätzlich die Möglichkeit, die Daten zu Zwecken der Archivierung zu speichern. Aufgrund der rechtlichen Anforderungen an die Ablage und Aufbewahrung der Daten wird diese Möglichkeit in dieser Arbeit nicht betrachtet. Vgl. hierzu Kapitel 2.4.3.2. Eine Archivierung findet ausschließlich auf unternehmensinternen Speichermedien der Offline-Ebene statt.

⁸³⁸ Vgl. Iannucci, Gupta / Cloud Data Center / 15; Pallickara, Pallickara, Pierce / Data Management in the Cloud / 528; Mahmood, Saeed / Frameworks for the Cloud / 314-315; Jewell et al. / Performance and Capacity / 10-11; Vgl. Kapitel 3.2

kann dabei einen beliebigen hohen Zugriffswert annehmen, dargestellt durch das „x“ im entsprechenden Intervall. Sinken die Nutzungsgrade, erfolgt die Zuordnung der Informationen zu den Klassen „Online“, „Nearline“ und „Offline“. ⁸³⁹

5.2.3.2 Implementierung der Verlagerungsfunktionen

Die in Kapitel 5.2.3.1 vorgestellten Klassifizierungsfunktionen dienen zur Klassifizierung der Daten anhand des Nutzungsgrads. Ändert sich der Nutzungsgrad und werden die Daten einer entsprechenden Klasse bzw. Speicherebene zugeordnet, ist eine Verlagerung bzw. Migration der Daten notwendig. ⁸⁴⁰ Um im späteren Verlauf der vorliegenden Arbeit im Rahmen der Wirtschaftlichkeitsanalyse den Kostenaufwand für Verlagerungsvorgänge berechnen zu können, ⁸⁴¹ muss die Anzahl der Verlagerungen für die verschiedenen Lebenszyklen und Simulationsläufe ⁸⁴² ermittelt werden. Hierzu werden durch Verwendung der WENN-Funktion in Microsoft Excel entsprechende Funktionen implementiert. Die Verlagerungsfunktion $V(K_n)$ ergibt sich wie folgt:

$$V(K_n) = (\text{WENN}(K_{n;t} < K_{n;t-1}; 1; 0))$$

Die Funktion $V(K_n)$ prüft, ob der untersuchte Klassenwert ⁸⁴³ (K_n) zum Zeitpunkt (t) ungleich des Klassenwerts (K_n) zum Zeitpunkt (t-1) ist. ⁸⁴⁴ Ist die Aussage wahr, liefert die Funktion den Wert „1“ zurück. In diesem Fall hat eine Verlagerung bzw. Migration

⁸³⁹ Vgl. Kapitel 5.2.3.1.1.1

⁸⁴⁰ Vgl. Kapitel 5.1.2.3, Kapitel 5.1.2.4

⁸⁴¹ Vgl. Kapitel 5.2.3.3.2.1.1.3, Kapitel 5.2.3.3.2.1.2.3, Kapitel 5.3

⁸⁴² Vgl. Kapitel 5.2.2.1.3.1

⁸⁴³ Die Werte der Klassen werden durch die Klassifizierungsfunktionen berechnet. Vgl. hierzu Kapitel 5.2.3.1.1.

⁸⁴⁴ Da die Klassifizierung basierend auf den Monatswerten der Nutzungsgrade erfolgt (Vgl. Kapitel 5.2.2.1.2), handelt es sich bei den Zeitpunkten um den jeweils betrachteten Monat (t) sowie den entsprechenden Vormonat (t-1). Die Lebenszyklen werden über einen Zeitraum von 5 Jahren betrachtet (Vgl. Kapitel 5.2.2.1.3). Im jeweils ersten Monat des jeweils ersten Jahres wird der Wert für den Verlagerungsvorgang im Rahmen der prototypischen Implementierung für jeden Simulationslauf und für jeden Lebenszyklus manuell auf „1“ gesetzt. Für den ersten Monat eines Lebenszyklus wird also angenommen, dass eine Verlagerung der Daten durchgeführt wird, da die Daten im Zuge der Implementierung eines ILM-Konzepts zunächst in die jeweilige Speicherebene, wie „Cloud“, „Nearline“ oder „Offline“, gelangen müssen. Es wird weiterhin angenommen, dass die Daten vor der Implementierung des ILM-Konzepts und der Durchführung der automatisierten Klassifizierung auf anderen Speichermedien gespeichert waren. Bezüglich der Online-Ebene wäre denkbar, dass die Daten vorher auf Speichermedien mit längeren Zugriffszeiten (Vgl. Kapitel 5.2.3.1.1.1) abgelegt waren und deshalb im jeweiligen ersten Monat zunächst auf die Online-Ebene verlagert werden müssen.

der Daten stattgefunden.⁸⁴⁵ Ist die Aussage falsch bzw. sind die beiden betrachteten Klassenwerte gleich, liefert die Funktion den Wert „0“ zurück.⁸⁴⁶ Die Funktion $V(K_n)$ ist für die beiden betrachteten Szenarien⁸⁴⁷ gleichermaßen anwendbar. Für das zweite Szenario werden zusätzlich Funktionen implementiert, die die Anzahl der Verlagerungsvorgänge zwischen unternehmensinternen Speichermedien⁸⁴⁸ und der betrachteten Cloud-Plattform⁸⁴⁹ ermitteln.⁸⁵⁰ Ein Verlagerungsvorgang, der beim Import der Daten in eine Cloud-Plattform entsteht, wird durch die Funktion $V_{CI}(K_n)$ ermittelt:

$$V_{CI}(K_n) = (\text{WENN}(\text{UND}((K_{n,t-1} > 4; K_{n,t} = 4; 1; 0))))$$

Mit Hilfe der Funktion $V_{CI}(K_n)$ wird für einen betrachteten Monat (t) überprüft, ob der untersuchte Klassenwert⁸⁵¹ (K_n) zum Zeitpunkt (t-1) bzw. im Vormonat ungleich des Klassenwerts (4) war und ob der Klassenwert im jeweils betrachteten Monat (t) den Wert (4) angenommen hat.⁸⁵² Sind beide Bedingungen wahr, findet eine Verlagerung der Daten in die Cloud-Ebene statt. Ein Verlagerungsvorgang, der beim Export der Daten aus einer Cloud-Plattform entsteht, wird durch die Funktion $V_{CE}(K_n)$ ermittelt:

$$V_{CE}(K_n) = (\text{WENN}(\text{UND}((K_{n,t-1} = 4; K_{n,t} < 4; 1; 0))))$$

Die Funktion $V_{CE}(K_n)$ liefert eine Aussage darüber, ob für einen Monat (t) der Klassenwert (K_n) zum Zeitpunkt (t-1) gleich des Klassenwerts (4) war und ob der Klassen-

⁸⁴⁵ Wenn bspw. der Wert der Klasse zum Zeitpunkt (t) gleich „2“ beträgt bzw. die Daten der Ebene „Nearline“ zugeordnet wurden, und der Wert der Klasse zum Zeitpunkt (t-1) bzw. im Vormonat noch den Wert „3“ bzw. „Online“ hatte, dann hat eine Verlagerung der Daten von der Online-Ebene zur Nearline-Ebene stattgefunden.

⁸⁴⁶ In diesem Fall wurden die Daten nicht verlagert.

⁸⁴⁷ Vgl. Kapitel 5.2.3.1.1.1

⁸⁴⁸ Vgl. Kapitel 5.1.1 und Kapitel 5.1.2.1

⁸⁴⁹ Vgl. Kapitel 5.2.3.3.2.1.2

⁸⁵⁰ Da Anbieter von Cloud-Diensten Gebühren für den Import der Daten in die Cloud-Plattform und den Export der Daten aus der Cloud-Plattform berechnen (Vgl. Kapitel 3.4.1), werden die Verlagerungsvorgänge für die in Kapitel 5.3 durchgeführte Wirtschaftlichkeitsanalyse ermittelt.

⁸⁵¹ Die Werte der Klassen werden durch die Klassifizierungsfunktionen berechnet. Vgl. hierzu Kapitel 5.2.3.1.1.

⁸⁵² Der Klassenwert 4 repräsentiert die Ebene „Cloud“. Vgl. Kapitel 5.2.3.1.1.2.

wert im jeweils betrachteten Monat (t) einen anderen Wert als (4) angenommen hat.⁸⁵³ Sind beide Bedingungen wahr, findet ein Export der Daten aus der Cloud-Ebene statt.⁸⁵⁴

Die vorgestellten und implementierten Verlagerungsfunktionen unterstellen, dass bei einer Abweichung der Klassenwerte innerhalb der betrachteten Monate eine Verlagerung stattgefunden hat. Dies resultiert aus der Tatsache, dass die implementierten Klassifizierungsfunktionen die Daten auf Basis des Nutzungsgrads klassifizieren und eine geeignete Speicherebene vorschlagen, auf die die Daten verlagert werden sollen. In der Praxis sollten die Vorschläge der Klassifizierungsfunktionen in jedem Fall durch Anwender und Administratoren kritisch überprüft werden, da die Verlagerung der Informationen in Abhängigkeit von der verwendeten Speicherebene mit Zeit- und Kostenaufwänden verbunden ist.⁸⁵⁵ Bei stark schwankenden Nutzungsgraden können bspw. häufige Verlagerungsvorgänge der Informationen zwischen den Speicherebenen notwendig werden.

5.2.3.3 Kostenfunktionen

5.2.3.3.1 Kostenarten und Kostenmodell

In dieser Arbeit wird eine Wirtschaftlichkeitsanalyse durchgeführt, um das Kosten- und Nutzenverhältnis der automatisierten Verwaltung unstrukturierter Daten im ILM zu untersuchen.⁸⁵⁶ Um die Kosten zu berechnen, die bei der Klassifizierung, Verlagerung und Speicherung der Daten entstehen, werden Kostenfunktionen implementiert.⁸⁵⁷ Die hierfür verwendeten Kostenarten⁸⁵⁸ und das zugrunde gelegte Kostenmodell werden im Folgenden vorgestellt. Die Kosten werden im Rahmen des Kostenmodells in einzelne Kostenarten gegliedert. Kostenmodelle eignen sich zur systematischen Darstellung und Gliederung von Kosten und Kostenarten.⁸⁵⁹ Mit Hilfe der Kostenfunktionen werden, basierend auf den Kostenarten, die Gesamtkosten für die in dieser Arbeit betrachteten

⁸⁵³ Der Klassenwert 4 repräsentiert die Ebene „Cloud“. Vgl. Kapitel 5.2.3.1.1.1.2.

⁸⁵⁴ Beispielsweise wenn die Daten der Cloud-Plattform zurück auf die unternehmensinternen Speichermedien übertragen werden.

⁸⁵⁵ Vgl. Kapitel 3.4.1

⁸⁵⁶ Vgl. Kapitel 5.3

⁸⁵⁷ Vgl. Kapitel 5.2.3.3.2

⁸⁵⁸ Kosten mit gleichen Merkmalen werden zu einer Kostenart zusammengefasst. Vgl. hierzu Ehrlenspiel, Kiewert, Lindemann / Kostenmanagement / 415; Schmidt / Kostenrechnung / 42-43.

⁸⁵⁹ Vgl. Charles, Schalk, Thiel / Kostenmodelle / 378; Liu, Ngoc, Laurent / Cost Framework / 4-5

Szenarien berechnet.⁸⁶⁰ Eine Kostenvergleichsrechnung ermöglicht später den Vergleich der Szenarien.

Ziel von ILM ist es, Informationen zu klassifizieren und kostengünstig zu verwalten.⁸⁶¹ Der Begriff Kosten wird definiert als der in Geldeinheiten bewertete sachzielbezogene Güterverzehr bzw. Ressourcenverbrauch innerhalb einer Periode.⁸⁶² Kosten werden in Kostenarten gegliedert.⁸⁶³ Die für die Klassifizierungsfunktionen relevanten Kostenarten werden in einem Kostenmodell zusammengefasst.⁸⁶⁴ Bezogen auf das ILM eignet sich die Gliederung der Kosten in Anschaffungskosten, Betriebskosten und sonstige Kosten.⁸⁶⁵ Bei den Anschaffungskosten handelt es sich um Aufwendungen bzw. Investitionen, die geleistet werden, um ein Wirtschaftsgut zu erwerben und in einen betriebsbereiten Zustand zu versetzen.⁸⁶⁶ Werden betriebliche Kapazitäten bereitgestellt und angeschafft, so handelt es sich um sogenannte Fixkosten, wenn diese unabhängig von der Beschäftigung oder Ausbringungsmenge anfallen.⁸⁶⁷ Zu den Anschaffungskosten (AK) zählen bspw. die Kosten für die Anschaffung von Speichermedien oder Archivierungssoftware.⁸⁶⁸ Bei den Betriebskosten handelt es sich um den Werteverzehr im laufenden Geschäftsbetrieb.⁸⁶⁹ Betriebskosten für den Speicherplatz (SBK)⁸⁷⁰ sind Kosten für die Miete⁸⁷¹ von Speichermedien bzw. Speicherplatz. Hierzu zählen ebenfalls die monatlich oder jährlich anfallenden Kosten für Cloud-Computing-Dienste bzw. IaaS. Die monatli-

⁸⁶⁰ Vgl. Kapitel 5.3 und Kapitel 5.2.3.3.2

⁸⁶¹ Vgl. Maier, Hädrich, Peinl / Knowledge Infrastructures / 254-257; Thome, Sollbach / Modelle des ILM / 22-30

⁸⁶² Vgl. Schmidt / Wirtschaftlichkeit / 262; Heinen / Führungslehre / 313

⁸⁶³ Vgl. hierzu VDI / Betriebswirtschaftliche Berechnungen / 5-6 ; Schröder / Operatives Controlling / 24.

⁸⁶⁴ Vgl. Tabelle 5-1

⁸⁶⁵ Vgl. hierzu Reichman, Whiteley, Chi / File Storage Costs / 5-7; Poulton / Data Storage Networking / 50-51. Die vorgeschlagene Kostenartengliederung und das in dieser Arbeit verwendete Kostenmodell resultieren aus der Literaturanalyse zum ILM und Storage-Management (Vgl. Kapitel 1.3).

⁸⁶⁶ Vgl. Bundesministerium der Justiz / HGB / §§ 255 Abs. 1; Scheffler / Steuerbilanz / 146-153

⁸⁶⁷ Vgl. Deimel, Heupel, Wiltner / Controlling / 241

⁸⁶⁸ Vgl. Kapitel 2.4.2; Kostenkategorie „Anschaffungskosten“ in Tabelle 5-1

⁸⁶⁹ Vgl. Röhrich / Investitionsrechnung / 13; Töpfer / Betriebswirtschaftslehre / 973

⁸⁷⁰ Vgl. hierzu Kostenart „Betriebskosten Speicherplatz“ in Tabelle 5-1.

⁸⁷¹ Hard- und Softwarekosten können sowohl den Anschaffungskosten als auch den Betriebskosten zugeordnet werden (Vgl. Tabelle 5-1). Werden bspw. Hardwarekomponenten gekauft, so stellen diese eine Investition dar und sind den Anschaffungskosten zuzuordnen. Vgl. hierzu Badach, Rieger / Netzwerk-Projekte / 152. Wird die Hardware oder Software über Dienstleister beschafft bzw. gemietet, so werden die Kosten den Betriebskosten zugeordnet. Vgl. hierzu Brugger / IT Business Case / 73.

chen Kosten für die Speichermedien oder IaaS-Dienste könnten ebenfalls den Mietkosten innerhalb der allgemeinen Betriebskosten zugeordnet werden.⁸⁷² Da die Speichermedien in den verschiedenen Speicherebenen für das ILM jedoch von hoher Relevanz sind, wird im Kostenmodell hierfür eine gesonderte Kostenart definiert.⁸⁷³ Administrationskosten (ABK)⁸⁷⁴ sind Betriebskosten für die automatisierte Klassifizierung im ILM. Dazu gehören Kosten für die Untersuchung der Metadaten⁸⁷⁵ und die Zuordnung unstrukturierter Daten zu Dokumentenklassen⁸⁷⁶, Kosten für die Anpassung und Pflege der Klassifizierungsfunktionen,⁸⁷⁷ Kosten für Zugriffe auf die Daten⁸⁷⁸ oder Kosten für die Verwaltung der Speichermedien und Speicherebenen⁸⁷⁹. Hierzu gehört auch gemietete Hard- und Software, wie bspw. Server-Hardware, um die Speichermedien betreiben zu können, oder Software zur Verwaltung der Daten in den verschiedenen Speicherebenen.⁸⁸⁰ Betriebskosten in Bezug auf die Verlagerung der Daten (VBK)⁸⁸¹ sind Kosten für die Verlagerung bzw. Migration der Daten⁸⁸² zwischen den verschiedenen Speicherebenen.⁸⁸³ Allgemeine Betriebskosten (GBK)⁸⁸⁴ sind Energiekosten⁸⁸⁵ oder Mietkosten⁸⁸⁶ für Büroräume⁸⁸⁷. Bei den sonstigen Kosten handelt es sich um Kosten, die einer

⁸⁷² Vgl. „Allgemeine Betriebskosten“ Tabelle 5-1

⁸⁷³ Vgl. Kostenart „Betriebskosten Speicherplatz“ in Tabelle 5-1

⁸⁷⁴ Vgl. hierzu Kostenart „Betriebskosten Administration“ in Tabelle 5-1.

⁸⁷⁵ Vgl. Kapitel 5.2.2.1.2

⁸⁷⁶ Vgl. Kapitel 5.2.2.1.3

⁸⁷⁷ Vgl. Kapitel 5.2.3.1.1; Die Klassifizierung von Informationen im ILM ist ohne automatisierte Unterstützung praktisch kaum möglich und wirtschaftlich nicht sinnvoll. Vgl. hierzu Bhagwan et al. / Management of Data Storage / ; Shah et al. / Classification for ILM / 1-15. Wenn sich bspw. die definierten Intervalle für die Speicherebenen aufgrund neuer Daten oder Speicherebenen ändern, so müssen die Klassifizierungsfunktionen entsprechend angepasst werden. Vgl. hierzu Kapitel 5.2.3.1.

⁸⁷⁸ Bei der Nutzung von Cloud-Diensten werden bspw. vom Anbieter der Dienste Gebühren für den Zugriff auf gespeicherte Dateien erhoben. Vgl. hierzu Kapitel 5.2.3.3.2.1.2.2.

⁸⁷⁹ Vgl. Kapitel 1.1; Kapitel 5.1.2.3

⁸⁸⁰ Vgl. Kapitel 2.4.2; Krcmar / Informationsmanagement / 258; Elmasri, Navathe / Datenbanksysteme / 369; Illik / Verteilte Systeme / 277

⁸⁸¹ Vgl. hierzu Kostenart „Betriebskosten Verlagerung Daten“ in Tabelle 5-1.

⁸⁸² Beispielsweise die Kosten, die bei der Verlagerung der Daten von der Online-Ebene zur Nearline-Ebene, entstehen.

⁸⁸³ Vgl. Kapitel 5.1.2.4

⁸⁸⁴ Vgl. hierzu Kostenart „Allgemeine Betriebskosten“ in Tabelle 5-1.

⁸⁸⁵ Zu den Energiekosten gehören bspw. Stromkosten. Vgl. hierzu Lampe / Green-IT / 59.

⁸⁸⁶ Vgl. Brugger / IT Business Case / 76

⁸⁸⁷ Vgl. Brugger / IT Business Case / 73

Kostenart nicht direkt zugeordnet sind, wie bspw. Werbekosten, Reisekosten, Kosten für Versicherungen oder Kosten für sonstige Beratungsdienstleistungen.⁸⁸⁸ Zu den sonstigen Kosten (SK) zählen bspw. Versicherungskosten⁸⁸⁹ oder Reisekosten.⁸⁹⁰

Kostenkategorie	Kostenart	Materialkosten	Personalkosten
Anschaffungskosten (AK)	Hardwarekosten	x	-
	Softwarekosten	x	-
	Schulungskosten	x	x
	Sonstige Anschaffungskosten	x	x
Betriebskosten (BK)	Betriebskosten Speicherplatz (SBK)		
	Kosten Speichermedien und Cloud-Speicher	x	-
	Betriebskosten Administration (ABK)		
	Kosten Erhebung Metadaten	x	x
	Kosten Zuordnung Daten zu Dokumentenklassen	x	x
	Kosten Anpassung und Pflege Klassifizierungsfunktionen	x	x
	Kosten Zugriffe auf Daten	x	x
	Kosten Verwaltung Speicherebenen	x	x
	Betriebskosten Verlagerung Daten (VBK)		
	Kosten Verlagerung- bzw. Migration Daten	x	x
	Allgemeine Betriebskosten (GBK)		
	Mietkosten	x	-
	Energiekosten	x	-
Sonstige Kosten (SK)	Sonstige Kosten für ILM	x	x

Tab. 5-1: Kostenmodell und Kostenarten⁸⁹¹

⁸⁸⁸ Vgl. Brugger / IT Business Case / 70; Gonschorrek, Hoffmeister / Planungs- und Entscheidungsprozesse / 502; Szyska / Kostenrechnung / 105

⁸⁸⁹ Wie beispielsweise die Elektronikversicherung. Vgl. hierzu Münch / Datenschutz / 189-190.

⁸⁹⁰ Vgl. Gonschorrek, Hoffmeister / Planungs- und Entscheidungsprozesse / 502

⁸⁹¹ Die gewählte Darstellung zur Aufteilung der Kosten in Kostenkategorien, Kostenarten sowie Material- und Personalkosten findet in Anlehnung an Wolff / IT-Controlling / 42 statt. Die Gesamtkosten werden durch die Addition von Anschaffungskosten, Betriebskosten und sonstigen Kosten berechnet. Vgl. Kapitel 5.3.

In Tabelle 5-1 ist dargestellt, welche Kostenarten Material- und Personalkosten beinhalten.⁸⁹² Bei zahlreichen Kostenarten sind Personalkosten enthalten. Dies hat einen hohen Anteil an Betriebskosten- und Administrationskosten, gemessen an den Gesamtkosten, zur Folge.⁸⁹³ Im Folgenden wird erörtert, welche Annahmen für die einzelnen Kostenkategorien und Kostenarten im Rahmen der durchgeführten Wirtschaftlichkeitsanalyse getroffen werden.⁸⁹⁴

5.2.3.3.1.1 Anschaffungskosten

Die Anschaffungskosten (AK) setzen sich aus Hardware-, Software- und Schulungskosten sowie sonstigen Anschaffungskosten zusammen.⁸⁹⁵ Für die in dieser Arbeit betrachteten Szenarien zur Wirtschaftlichkeitsanalyse im ILM werden bezüglich der Anschaffungskosten folgende Annahmen getroffen:

- Es fallen keine Anschaffungs- bzw. Investitionskosten für Hard- und Software an. Die Kosten für gemietete Hardware und Software, die bspw. für den Betrieb bzw. die Verwaltung der Speichermedien benötigt wird,⁸⁹⁶ sind Bestandteil der Betriebskosten für die Administration.⁸⁹⁷ Die für das ILM verantwortlichen Mitarbeiter müssen geschult werden. Administratoren und IT-Verantwortliche erhalten eine Schulung zu bspw. den Inhalten Konzept des ILM, Vorgehensweise bei der Vorklassifizierung der unstrukturierten Daten, Auswahl von Werkzeugen zur Untersuchung der Metadaten, Verwaltung der Daten in unterschiedlichen Speicherebenen und rechtliche Anforderungen an die Verwaltung der Daten.⁸⁹⁸ Die hierfür angenommene Schulungsdauer und die Schulungskosten sind in Anhang 5.1 enthalten.

⁸⁹² Vgl. Anhang 5.1 und Anhang 5.3

⁸⁹³ Vgl. Kapitel 5.2.3.3.1.2

⁸⁹⁴ Dies bildet die Grundlage für die nachfolgende Implementierung der Kostenfunktionen. Vgl. hierzu Kapitel 5.2.3.3.2.

⁸⁹⁵ Vgl. Kostenkategorie „Anschaffungskosten“ Tabelle 5-1

⁸⁹⁶ Vgl. Kapitel 5.2.3.3.1

⁸⁹⁷ Vgl. „Kosten Verwaltung Speicherebenen“ Tabelle 5-1; vgl. hierzu die Annahmen zu den Betriebskosten Kapitel 5.2.3.3.1.2

⁸⁹⁸ Vgl. Short / ILM / 21-22

- Bei der Entwicklung und Implementierung eines ILM-Konzepts zur Verwaltung der Daten auf unterschiedlichen Speicherebenen, der Implementierung der Funktionen zur automatisierten Klassifizierung unstrukturierter Daten und der Untersuchung der Daten hinsichtlich unternehmensinterner und rechtlicher Anforderungen entstehen sonstige Anschaffungskosten. Die hierfür angenommenen Aufwände sind in Anhang 5.1 enthalten. Zur Berechnung der Anschaffungskosten ist es nicht notwendig, Kostenfunktionen zu implementieren. Es wird angenommen, dass die Anschaffungskosten, bezogen auf die untersuchten Lebenszyklen und den betrachteten Zeitraum von fünf Jahren, jeweils im ersten Jahr anfallen und unabhängig von den Klassifizierungsergebnissen und Simulationsläufen entstehen. Die Anschaffungskosten werden als Zellenwerte in Microsoft Excel hinterlegt und bei der Berechnung der Kosten und der Durchführung der Kostenvergleichsrechnung zu den Betriebskosten hinzuaddiert.⁸⁹⁹ Die Betriebskosten werden mit Hilfe von Kostenfunktionen berechnet, da diese verbrauchsabhängig sind und auf den Klassifizierungsergebnissen basieren.⁹⁰⁰

5.2.3.3.1.2 Betriebskosten

Im ILM sind die Betriebs- und Administrationskosten von besonderer Bedeutung. Die Kosten für den Betrieb bzw. die Administration der Speichermedien⁹⁰¹ sind um das vier- bis achtfache höher als die Kosten der Datenträger.⁹⁰² Der Grund hierfür ist, dass in zahlreichen Kostenarten, die den Betriebskosten zugehörig sind, ein Personalkostenanteil enthalten ist.⁹⁰³ Personalkosten stellen typischerweise einen hohen Anteil, bezogen auf die Gesamtkosten einer IT-Investition, dar.⁹⁰⁴ Das ist auch bei ILM der Fall.⁹⁰⁵ Trotz der angestrebten Automatisierung im ILM werden Anwender und Administratoren bspw. bei der Klassifizierung der Daten und der Verwaltung der Speichermedien

⁸⁹⁹ Vgl. Kapitel 5.3

⁹⁰⁰ Vgl. Kapitel 5.2.3.3.2

⁹⁰¹ Vgl. Kapitel 5.1.2.3

⁹⁰² Vgl. hierzu Glazer / Measuring Information / 99-110; Kaplan, Roy, Srinivasaraghavan / Demand data storage / 1; Forrest, Kaplan, Kindler / Data centers / 1-10; Forrest, Kaplan, Kindler / Energy Efficiency / 1-13; Gantz, Reinsel / Extracting Value / 4; Schadler / Email in the Cloud / 4.

⁹⁰³ Vgl. Tabelle 5-1

⁹⁰⁴ Vgl. Zsifkovits / Projektsteuerung / 323

⁹⁰⁵ Vgl. BITKOM / Leitfaden ILM / 6; Short / ILM / 21-22

benötigt.⁹⁰⁶ Für die in dieser Arbeit betrachteten Szenarien⁹⁰⁷ zur Wirtschaftlichkeitsanalyse im ILM werden bezüglich der Betriebskosten folgende Annahmen getroffen:

- Dienstleister berechnen die Speicherplatzkosten (SBK)⁹⁰⁸ auf Basis des Durchschnitts aus verschiedenen Tagen eines Rechnungsmonats.⁹⁰⁹ Die hierfür anfallenden Kosten können direkt den Betriebsausgaben zugerechnet werden, eine Abschreibung ist nicht erforderlich.⁹¹⁰ Die Administrationskosten im ILM, bestehend aus den Betriebskosten für die Administration (ABK)⁹¹¹ und den allgemeinen Betriebskosten (GBK)⁹¹², werden im Rahmen der Wirtschaftlichkeitsanalyse und der durchgeführten Simulation mit Hilfe eines Betriebskostenfaktors (BKF) berechnet.⁹¹³ Da die Wirtschaftlichkeitsanalyse zum ILM in dieser Arbeit nicht bezogen auf eine einzelne Organisation durchgeführt wird und die Betriebskosten nicht aus einem einzelnen Unternehmen erhoben werden, werden die Administrationskosten (ABK) und die allgemeinen Betriebskosten (GBK) durch Verwendung des Betriebskostenfaktors (BKF) berechnet. Der Faktor drückt aus, dass die Kosten für den Betrieb von ILM bzw. für die Speichermedien und Speicherebenen um ein Vielfaches höher sind, als die Kosten für die Speichermedien bzw. den Speicherplatz.⁹¹⁴ Für die Verwaltung der verschiedenen Speicherebenen werden unterschiedliche Betriebskostenfaktoren zugrunde gelegt. Im Hinblick auf die zwei betrachteten Szenarien wird bspw. angenommen, dass die Verwaltung unstrukturierter Daten auf unternehmensinternen

⁹⁰⁶ Vgl. Kapitel 5.1.3.2

⁹⁰⁷ Vgl. Kapitel 5.2.3.1.1.1

⁹⁰⁸ Vgl. hierzu Kostenart „Betriebskosten Speicherplatz“ in Tabelle 5-1.

⁹⁰⁹ Vgl. Turczyk / Wertzuweisung / 462; Laudon, Laudon, Schoder / Wirtschaftsinformatik / 366

⁹¹⁰ Vgl. Stollenwerk / Wertschöpfungsmanagement / 173; Goldstein / Kontieren / 284-285

⁹¹¹ Vgl. hierzu Kostenart „Betriebskosten Administration“ in Tabelle 5-1. Für das zweite Szenario werden zusätzlich die Kosten für Zugriffe auf die Daten der Cloud-Ebene betrachtet, da die Nutzung der Daten bzw. die Zugriffe auf die Daten vom Cloud-Anbieter gesondert abgerechnet werden. Vgl. hierzu Kapitel 5.2.3.3.2.1.2.2.

⁹¹² Vgl. hierzu Kostenart „Allgemeine Betriebskosten“ in Tabelle 5-1.

⁹¹³ Vgl. zur Implementierung der Kostenfunktionen Kapitel 5.2.3.3.2.

⁹¹⁴ Vgl. hierzu Glazer / Measuring Information / 99-110; Kaplan, Roy, Srinivasaraghavan / Demand data storage / 1; Forrest, Kaplan, Kindler / Data centers / 1-10; Forrest, Kaplan, Kindler / Energy Efficiency / 1-13; Gantz, Reinsel / Extracting Value / 4; Schadler / Email in the Cloud / 4. Vgl. hierzu Anhang 5.3 zur Beschreibung der implementierten Betriebskostenfaktoren innerhalb der Kostenfunktionen und zur Berechnung der Betriebskosten.

Speichermedien einen höheren Aufwand verursacht, als die Verwaltung der Daten mit Hilfe von Cloud-Diensten.⁹¹⁵ Weiterhin wird angenommen, dass die für die Verwaltung bzw. den Betrieb der Speichermedien gemietete Hardware⁹¹⁶ ausschließlich für die Verwaltung der betrachteten Daten⁹¹⁷ genutzt wird.⁹¹⁸

- Es wird für beide Szenarien angenommen, dass die Speicherplatzkosten (SBK), bezogen auf den betrachteten Zeitraum von fünf Jahren, jährlich um dreißig Prozent⁹¹⁹ sinken.⁹²⁰
- Die Werte⁹²¹ der jeweiligen Betriebskostenfaktoren (BKF) der unterschiedlichen Speicherebenen steigen, über den betrachteten Zeitraum von fünf Jahren hinweg, jährlich um fünf Prozent.⁹²² Obwohl die Preise für Speicherplatz jährlich sinken und sich Speichermedien hinsichtlich Speicherkapazität und Übertragungsgeschwindigkeiten weiterentwickeln, steigt der Aufwand für die Administration bzw. Verwaltung der Speichermedien und -ebenen.⁹²³ Die Gründe hierfür sind die jährlich wachsende Datenmenge⁹²⁴ von 30% sowie steigende Lohn- und Energiekosten.⁹²⁵

⁹¹⁵ Vgl. zur Verwendung der Betriebskostenfaktoren zur Berechnung der Betriebskosten mit Hilfe der Kostenfunktionen Kapitel 5.2.3.3.2 und den zugrunde gelegten Betriebskostenfaktoren Anhang 5.3.

⁹¹⁶ Vgl. hierzu die Betriebskosten Administration (ABK) Kapitel 5.2.3.3.1

⁹¹⁷ Vgl. Kapitel 5.2.2.1

⁹¹⁸ Der Betrieb der Speichermedien der Ebenen Online, Nearline und Offline findet auf unterschiedlichen Hardwareressourcen statt. Vgl. hierzu Krcmar / Informationsmanagement / 258; Elmasri, Navathe / Datenbanksysteme / 369; Illik / Verteilte Systeme / 277. Werden bspw. die Daten von der Online-Ebene zur Nearline-Ebene verlagert, so gilt die Annahme, dass für die Online-Ebene solange keine Betriebskosten anfallen bzw. die Speichermedien und die Hardware dieser Ebene solange nicht genutzt werden, bis die Daten wieder zur Ebene Online zurückverlagert werden. Die Miete des Speicherplatzes und der benötigten Hardware findet auf monatlicher Basis statt. Vgl. hierzu van Husen / IT-Dienstleistungen / 64; Laudon, Laudon, Schoder / Wirtschaftsinformatik / 269-270; McLeod, Jordan / Systems Development / 465.

⁹¹⁹ Vgl. Gilheany / Decline Storage / 1-2; Jonscher / Economic Study IT / 21-22; Frenzel, Frenzel / Information Technology / 127-128; Turczyk et al. / Model ILM / 528

⁹²⁰ Vgl. Anhang 5-3

⁹²¹ Vgl. Anhang 5-3

⁹²² Vgl. Anhang 5-3

⁹²³ Vgl. Krcmar / Informationsmanagement / 258

⁹²⁴ Vgl. Kapitel 5.2.2.1

⁹²⁵ Vgl. Lippold / Unternehmensberatung / 46-47; Elias-Linde / Humanressourcenmanagement / 49-50; Erek, Schmidt, Schilling / Green-IT / 7; Liedloff, Bromberger / Datenhaltung / 571; Aichele / Smart Energy / 120

5.2.3.3.1.3 Sonstige Kosten

Für die Szenarien zur Wirtschaftlichkeitsanalyse wird angenommen, dass keine sonstigen Kosten (SN)⁹²⁶ anfallen. Im Mittelpunkt der Betrachtungen stehen die Anschaffungs- und Betriebskosten für die automatisierte Verwaltung unstrukturierter Daten im ILM.⁹²⁷

5.2.3.3.2 Implementierung der Kostenfunktionen

Die Berechnung der Kosten für die automatisierte Verwaltung unstrukturierter Daten im ILM erfolgt durch die Implementierung und Anwendung von Kostenfunktionen in Microsoft Excel.⁹²⁸ Entsprechend der zugeordneten Klassen,⁹²⁹ den genutzten Speichermedien und -ebenen⁹³⁰ sowie der Anzahl der Verlagerungsvorgänge⁹³¹, werden die Betriebskosten für die betrachteten Szenarien⁹³² ermittelt.⁹³³ Der Aufbau und die Funktionsweise der Kostenfunktionen werden in diesem Kapitel vorgestellt. In Anhang 5.2 befindet sich eine Übersicht zu den Parametern, die in den Kostenfunktionen verwendet werden.

5.2.3.3.2.1 Szenarien zur Klassifizierung unstrukturierter Daten

5.2.3.3.2.1.1 Szenario 1: Klassifizierung der Daten ohne Cloud-Ebene

5.2.3.3.2.1.1.1 Kostenfunktion zur Berechnung der Speicherplatzkosten

In Kapitel 5.2.3.3.1 wurden die Kostenarten vorgestellt, die für die Wirtschaftlichkeitsanalyse in dieser Arbeit relevant sind. Die Funktion zur Berechnung der Betriebskosten für den Speicherplatz (SBK₁) ergibt sich im ersten Szenario wie folgt:

⁹²⁶ Vgl. hierzu Kostenkategorie „Sonstige Kosten“ in Tabelle 5-1.

⁹²⁷ Vgl. Kapitel 5.2.3.3.1.1; Kapitel 5.2.3.3.1.2

⁹²⁸ Vgl. Kapitel 5.2.3.3.2.1; Hierzu wird die WENN-Funktion von Microsoft Excel verwendet.

⁹²⁹ Vgl. Kapitel 5.2.3.1.1

⁹³⁰ Vgl. Kapitel 5.2.3.1.1.1.1

⁹³¹ Vgl. Kapitel 5.2.3.2

⁹³² Vgl. Kapitel 5.2.3.1.1.1

⁹³³ Die ermittelten Betriebskosten dienen der in Kapitel 5.3 durchgeführten Kostenvergleichsrechnung zur Analyse der Wirtschaftlichkeit. Die Kostenfunktionen werden, bezogen auf die betrachteten Lebenszyklen und Szenarien, für jeden einzelnen Klassifizierungslauf, im Rahmen der Simulation, ausgeführt. Vgl. hierzu Kapitel 5.2.2.1.3. Bei den Lebenszyklen mit Rauschgraden werden für die jeweiligen Betriebskosten für den Speicherplatz, die Administration und Verlagerung der Daten, aus den 30 Simulationsläufen, die Mittelwerte berechnet.

$$\text{SBK}_1(\mathbf{K}_n) = (\text{WENN}(\mathbf{K}_{1n;t} = 3; \text{SBK}_{\text{Online};t} \cdot m_t; (\text{WENN}(\mathbf{K}_{1n;t} = 2; \text{SBK}_{\text{Nearline}} \cdot m_t; \text{SBK}_{\text{Offline}} \cdot m_t))))$$

Die Funktion $\text{SBK}_1(\mathbf{K}_n)$ prüft, welcher Klasse⁹³⁴ die Daten zum Zeitpunkt (t) zugeordnet sind. Entsprechend der Klasse und den damit verbundenen Speicherebenen werden die Speicherplatzkosten berechnet. Zunächst wird überprüft, ob die Daten der Klasse mit dem Wert „3“ zugeordnet wurden.⁹³⁵ Ist die Aussage wahr, werden die Kosten berechnet, die bei Verwendung der Ebene „Online“ entstehen. Hierzu werden die Betriebskosten für den genutzten Speicherplatz der Online-Ebene ($\text{SBK}_{\text{Online};t}$) mit der zum Zeitpunkt (t) vorhandenen Datenmenge (m) multipliziert.⁹³⁶ Ist die Aussage falsch, wird überprüft, ob die Daten der Klasse mit dem Wert „2“ zugeordnet wurden. Trifft dies zu, werden die Speicherplatzkosten entsprechend für die Ebene „Nearline“ berechnet. Trifft dies nicht zu, gehören die Daten der Ebene „Offline“ an und die Speicherplatzkosten werden für diese Ebene berechnet.

5.2.3.3.2.1.1.2 Kostenfunktion zur Berechnung der Administrationskosten

Die Funktion $\text{ABK}_1(\mathbf{K}_n)$ berechnet die Betriebskosten für die Administration der Speicherebenen⁹³⁷ im ILM:

$$\text{ABK}_1(\mathbf{K}_n) = (\text{WENN}(\mathbf{K}_{1n;t} = 3; (\text{SBK}_{\text{Online};t} \cdot \text{BKF}_{S;t} \cdot m_t); (\text{WENN}(\mathbf{K}_{1n;t} = 2; (\text{SBK}_{\text{Nearline};t} \cdot \text{BKF}_{S;t} \cdot m_t); \text{SBK}_{\text{Offline};t} \cdot \text{BKF}_{S;t} \cdot m_t))))$$

Sind die Daten der Klasse mit dem Wert „3“ bzw. der Ebene „Online“ zugeordnet, werden die Administrationskosten berechnet, indem die Speicherplatzkosten der Online-Ebene ($\text{SBK}_{\text{Online};t}$) mit dem Betriebskostenfaktor⁹³⁸ für Speichermedien ($\text{BKF}_{S;t}$) multipliziert werden. Das daraus resultierende Ergebnis wird wiederum mit der Datenmen-

⁹³⁴ Die Werte der Klassen ($\mathbf{K}_{1n;t}$) werden durch die Klassifizierungsfunktionen des ersten Szenarios berechnet. Vgl. hierzu Kapitel 5.2.3.1.1.1.1.

⁹³⁵ Die Klasse mit dem Wert „3“ entspricht der Ebene „Online“.

⁹³⁶ Die für den Speicherplatz zugrunde gelegten Betriebskosten, für den betrachteten Zeitraum von fünf Jahren, sind in Anhang 5.3 dargestellt. Vgl. Anhang 5.1 zu den relevanten Datenmengen und dem angenommenen Datenwachstum.

⁹³⁷ Vgl. Kapitel 5.2.3.3.1

⁹³⁸ Vgl. Anhang 5.3; Kapitel 5.2.3.3.1.2

ge (m_t) multipliziert.⁹³⁹ Die Berechnung der Administrationskosten für die Ebenen „Nearline“ und „Offline“ erfolgt analog.⁹⁴⁰

5.2.3.3.2.1.1.3 Kostenfunktion zur Berechnung der Verlagerungskosten

Die Funktion $VBK_1(K_n)$ berechnet die Betriebskosten für die Verlagerung⁹⁴¹ der Daten:

$$\begin{aligned} VBK_1(K_n) = & (\text{WENN}(K_{1n;t} = 3; (SBK_{\text{Online};t} \cdot BKF_{V;t} \cdot m_t) \cdot V(K_{n;t}); \\ & (\text{WENN}(K_{1n;t} = 2; (SBK_{\text{Nearline};t} \cdot BKF_{V;t} \cdot m_t) \cdot V(K_{n;t}); \\ & (SBK_{\text{Offline};t} \cdot BKF_{V;t} \cdot m_t) \cdot V(K_{n;t})))) \end{aligned}$$

Wurden die Daten der Ebene „Online“ zugeordnet, werden die Speicherplatzkosten der Online-Ebene ($SBK_{\text{Online};t}$) mit dem Betriebskostenfaktor für Verlagerungsvorgänge ($BKF_{V;t}$) und der Datenmenge (m_t) multipliziert. Das resultierende Ergebnis wird zusätzlich mit dem Verlagerungswert ($V(K_{n;t})$) multipliziert.⁹⁴² Die Berechnung der Verlagerungskosten für die Ebenen „Nearline“ und „Offline“ findet auf die gleiche Weise statt, unter Verwendung der Speicherplatzkosten der Nearline-Ebene ($SBK_{\text{Nearline};t}$) und Offline-Ebene ($SBK_{\text{Offline};t}$).⁹⁴³

5.2.3.3.2.1.2 Szenario 2: Klassifizierung der Daten mit Cloud-Ebene

5.2.3.3.2.1.2.1 Kostenfunktion zur Berechnung der Speicherplatzkosten

Die Funktion $SBK_2(K_n)$ berechnet im zweiten Szenario die Betriebskosten für den Speicherplatz:

⁹³⁹ Je größer die Datenmenge ist, desto höher ist der Aufwand zur Verwaltung der entsprechenden Speicherebenen.

⁹⁴⁰ Für den Fall, dass die Daten bei der Klassifizierung der Ebene „Nearline“ oder „Offline“ zugeordnet wurden.

⁹⁴¹ Vgl. Kapitel 5.2.3.3.1

⁹⁴² Vgl. Kapitel 5.2.3.2 zur Beschreibung der Verlagerungsfunktion und zur Berechnung des Verlagerungswerts. Der Verlagerungswert hat den Wert „1“, wenn im Vormonat eine Verlagerung der Daten, ausgehend von einer anderen Speicherebene, stattgefunden hat. Es werden also mit der Funktion $VBK(K_n)$ nur Verlagerungskosten für die zu prüfende Speicherebene berechnet, wenn die Daten im Vormonat ($t-1$) einer anderen Speicherebene, die ungleich der zu prüfenden Speicherebene im Monat (t) ist, zugeordnet waren.

⁹⁴³ Die jeweiligen Speicherplatzkosten werden ebenfalls mit dem Betriebskostenfaktor für Verlagerungsvorgänge, der Datenmenge und dem Verlagerungswert multipliziert.

$$\begin{aligned} \text{SBK}_2(\mathbf{K}_n) = & (\text{WENN}(\mathbf{K}_{2n;t} = 4; \text{SBK}_{\text{Cloud};t} \cdot m_t; (\text{WENN}(\mathbf{K}_{2n;t} = 3; \\ & \text{SBK}_{\text{Online}} \cdot m_t; (\text{WENN}(\mathbf{K}_{2n;t} = 2; \text{SBK}_{\text{Nearline}} \cdot m_t; \\ & \text{SBK}_{\text{Offline}} \cdot m_t)))))) \end{aligned}$$

Die Funktion $\text{SBK}_2(\mathbf{K}_n)$ prüft, ob die Daten der Klasse⁹⁴⁴ „Cloud“, „Online“, „Nearline“ oder „Offline“ zugeordnet sind. Entsprechend der Klasse und den damit verbundenen Speicherebenen werden die Speicherplatzkosten berechnet. Zunächst wird überprüft, ob die Daten der Klasse mit dem Wert „4“ zugeordnet wurden.⁹⁴⁵ Ist die Aussage wahr, werden die Kosten berechnet, die bei Verwendung von IaaS⁹⁴⁶ bzw. der Ebene „Cloud“ entstehen. Hierzu werden die Betriebskosten für den aktuell genutzten Speicherplatz der Cloud-Ebene ($\text{SBK}_{\text{Cloud};t}$) mit der zum Zeitpunkt (t) vorhandenen Datenmenge (m) multipliziert.⁹⁴⁷ Ist die Aussage falsch, wird überprüft, ob die Daten der Klasse mit dem Wert „3“ zugeordnet wurden. Trifft dies zu, werden die Speicherplatzkosten entsprechend für die Ebene „Online“ berechnet. Trifft dies nicht zu, gehören die Daten entweder der Ebene „Nearline“ oder der Ebene „Offline“ an und die Speicherplatzkosten werden für diese Ebenen entsprechend berechnet.

5.2.3.3.2.1.2.2 Kostenfunktion zur Berechnung der Administrationskosten

Die Funktion $\text{ABK}_2(\mathbf{K}_n)$ zur Berechnung der Betriebskosten für die Administration der Speicherebenen⁹⁴⁸ ergibt sich für das zweite Szenario wie folgt:

$$\begin{aligned} \text{ABK}_2(\mathbf{K}_n) = & (\text{WENN}(\mathbf{K}_{2n;t} = 4; (\text{SBK}_{\text{Cloud};t} \cdot \text{BKF}_{C;t} \cdot m_t) \\ & + (\text{ABK}_Z \cdot (n_t / 10.000)); (\text{WENN}(\mathbf{K}_{2n;t} = 3; \\ & (\text{SBK}_{\text{Online};t} \cdot \text{BKF}_{S;t} \cdot m_t); (\text{WENN}(\mathbf{K}_{2n;t} = 2; \\ & (\text{SBK}_{\text{Nearline};t} \cdot \text{BKF}_{S;t} \cdot m_t); (\text{SBK}_{\text{Offline};t} \cdot \text{BKF}_{S;t} \cdot m_t)))))) \end{aligned}$$

⁹⁴⁴ Die Werte der Klassen ($\mathbf{K}_{2n;t}$) werden durch die Klassifizierungsfunktionen des zweiten Szenarios berechnet. Vgl. hierzu Kapitel 5.2.3.1.1.1.2.

⁹⁴⁵ Die Klasse mit dem Wert „4“ entspricht der Ebene „Cloud“. Vgl. hierzu Kapitel 5.2.3.1.1.1.2.

⁹⁴⁶ Vgl. Kapitel 3.2

⁹⁴⁷ Die für den Speicherplatz zugrunde gelegten Betriebskosten, für den betrachteten Zeitraum von fünf Jahren, sind in Anhang 5.3 dargestellt. Vgl. Anhang 5.1 zu den relevanten Datenmengen und dem angenommenen Datenwachstum.

⁹⁴⁸ Vgl. Kapitel 5.2.3.3.1

Sind die Daten der Klasse mit dem Wert „4“ bzw. der Ebene „Cloud“ zugeordnet, werden die Administrationskosten berechnet, indem die Speicherplatzkosten der Cloud-Ebene ($SBK_{Cloud; i}$) mit dem Betriebskostenfaktor für Cloud-Speicher ($BKF_{C; i}$) und der Datenmenge (m_i) multipliziert werden. Zum daraus resultierenden Wert werden die Kosten für Zugriffe⁹⁴⁹ auf die Daten der Ebene Cloud addiert. In dieser Arbeit wird angenommen, dass zur Verwaltung der unstrukturierten Daten in der Cloud der Simple Storage Service (Amazon S3) vom Anbieter Amazon verwendet wird. Amazon S3 ist der derzeit bekannteste und am weitesten verbreitete Dienst für Cloud-Speicher.⁹⁵⁰ Amazon berechnet Anforderungsgebühren pro 10.000 Zugriffe.⁹⁵¹ Gehören die Daten der Ebene „Online“, „Nearline“ oder „Offline“ an, werden die Administrationskosten, analog zur Vorgehensweise zur Berechnung der Administrationskosten im ersten Szenario, ermittelt.⁹⁵²

5.2.3.3.2.1.2.3 Kostenfunktion zur Berechnung der Verlagerungskosten

Die Wirtschaftlichkeitsanalyse für die verschiedenen Lebenszyklen der Informationen erfolgt mit drei verschiedenen Datenmengen.⁹⁵³ Für das jeweils erste Jahr, bezogen auf den betrachteten Zeitraum von fünf Jahren, werden Datenmengen von 1 Terabyte, 10 Terabyte und 100 Terabyte betrachtet.⁹⁵⁴ Die Funktion $VBK_2(K_n)$ berechnet die Betriebskosten für die Verlagerung⁹⁵⁵ der Daten für den ersten Fall, der Datenmenge von 1 Terabyte im ersten Jahr:⁹⁵⁶

⁹⁴⁹ Die Kosten für die Zugriffe auf die Daten in der Cloud werden durch den Parameter (ABK_Z) repräsentiert. Vgl. hierzu Kapitel 5.2.3.3.1.

⁹⁵⁰ Vgl. Magoulès, Pan, Teng / Data Intensive Computing / 12; Darcey, Conder / Android Application Programming / 227; Buyya, Vecchiola, Selvi / Cloud Computing / 92; Cusumano / SaaS and Cloud Computing / 10

⁹⁵¹ Vgl. Amazon S3 / S3-Preise / und Anhang 5.3 für die berechneten Gebühren.

⁹⁵² Vgl. Kapitel 5.2.3.3.2.1.1.2

⁹⁵³ Vgl. hierzu Kapitel 5.2.2.1

⁹⁵⁴ Vgl. hierzu Anhang 1

⁹⁵⁵ Vgl. Kapitel 5.2.3.3.1

⁹⁵⁶ Die Datenmenge beträgt im ersten Jahr 1 Terabyte und das Datenwachstum beträgt 30% pro Jahr, für die darauf folgenden vier Jahre. Vgl. hierzu Kapitel 5.2.2.1, Anhang 1.

$$\begin{aligned} \text{VBK}_2(\mathbf{K}_n) = & (\text{WENN}(\mathbf{K}_{2n;t} = 4; (\text{VBK}_{\text{Im-Cloud};t} \cdot (a_t / 1.000) \cdot V_{\text{CI}}(\mathbf{K}_{n;t})); \\ & (\text{WENN}(\mathbf{K}_{2n;t} = 3; ((\text{SBK}_{\text{Online};t} \cdot \text{BKF}_{V;t} \cdot m_t) \cdot V(\mathbf{K}_{n;t})) + \\ & (\text{VBK}_{\text{Ex-Cloud};t} \cdot V_{\text{CE}}(\mathbf{K}_{n;t}) \cdot m_t); \\ & (\text{WENN}(\mathbf{K}_{2n;t} = 2; ((\text{SBK}_{\text{Nearline};t} \cdot \text{BKF}_{V;t} \cdot m_t) \cdot V(\mathbf{K}_{n;t})) + \\ & (\text{VBK}_{\text{Ex-Cloud};t} \cdot V_{\text{CE}}(\mathbf{K}_{n;t}) \cdot m_t); \\ & ((\text{SBK}_{\text{Offline};t} \cdot \text{BKF}_{V;t} \cdot m_t) \cdot V(\mathbf{K}_{n;t})) + (\text{VBK}_{\text{Ex-Cloud};t} \cdot \\ & V_{\text{CE}}(\mathbf{K}_{n;t}) \cdot m_t)))))) \end{aligned}$$

In der Funktion $\text{VBK}_2(\mathbf{K}_n)$ stellen die Verlagerungsvorgänge, bezogen auf den Import in die Cloud-Ebene und den Export aus der Cloud-Ebene⁹⁵⁷, eine Besonderheit dar. Erfolgte die Zuordnung der Daten zur Ebene „Cloud“ bzw. zur Klasse mit dem Wert „4“, berechnet die Funktion die Kosten für den Import der Daten in die Cloud-Ebene.⁹⁵⁸ Hierzu werden die Anforderungskosten ($\text{VBK}_{\text{Im-Cloud};t}$) berechnet⁹⁵⁹ und mit dem Verlagerungswert für den Cloud-Import⁹⁶⁰ ($V_{\text{CI}}(\mathbf{K}_{n;t})$) multipliziert.⁹⁶¹ Über die sogenannte „PUT-Anforderung“ können in Amazon S3 Dateien angelegt werden. Amazon berechnet die Gebühr für das Anlegen von neuen Dateien pro 1.000 Dateien.⁹⁶² In der Wirtschaftlichkeitsanalyse und den Simulationsläufen wird die Anzahl der verwalteten Dateien nicht betrachtet.⁹⁶³ Da jedoch zusätzlich die Annahme gilt, dass die Daten im zweiten Szenario, für die Datenmenge von 1 Terabyte im ersten Jahr und einem jährli-

⁹⁵⁷ Vgl. Kapitel 5.2.3.2

⁹⁵⁸ Es wird angenommen, dass die Datenmenge, die im ersten Jahr 1 Terabyte beträgt, eingeschlossen dem Datenwachstum von 30% in den jeweiligen Folgejahren (Vgl. Kapitel 5.2.2.1, Anhang 1), über das Internet in die Cloud-Ebene übertragen wird. Amazon berechnet zwar keine Datenübertragungsgebühren, wenn die Daten über das Internet in die Cloud-Ebene übertragen bzw. importiert werden. Vgl. hierzu Amazon S3 / S3-Preise / . Allerdings berechnet Amazon sogenannte Anforderungsgebühren, da das Erstellen von einzelnen Dateien kostenpflichtig ist.

⁹⁵⁹ Die Anforderungskosten pro 1.000 Dateien werden berechnet, indem die Anforderungsgebühr ($\text{VBK}_{\text{Im-Cloud};t}$) mit der Anzahl der Dateien (a_t) multipliziert wird. Die Anzahl der Dateien wird zuvor durch den Wert 1.000 geteilt.

⁹⁶⁰ Vgl. Kapitel 5.2.3.2; Mit Hilfe des Verlagerungswerts des Cloud-Imports wird überprüft, ob die Daten im Vormonat ($t-1$) auch bereits der Cloud-Ebene angehört haben. Ist dies der Fall, hat kein Clou-Import im Monat (t) stattgefunden, der Verlagerungswert hätte den Wert „0“.

⁹⁶¹ Die Funktion muss nicht gesondert prüfen, ob die Daten aus der Cloud-Ebene exportiert wurden. Es kann ausgeschlossen werden, dass ein Export der Daten aus der Cloud-Ebene im Monat ($t-1$) stattgefunden hat, wenn die Daten der Cloud-Ebene im Monat (t) zugeordnet wurden. Eine Verlagerung der Daten findet ausschließlich zwischen unterschiedlichen Speicherebenen statt. Aus diesem Grund muss die Funktion $\text{VBK}_2(\mathbf{K}_n)$ nicht prüfen, ob ein Export der Daten aus der Cloud-Ebene stattgefunden hat, wenn sich die Daten bereits in der Cloud-Ebene befinden.

⁹⁶² Vgl. Amazon S3 / S3-Preise /

⁹⁶³ Vgl. Kapitel 5.2.2.1 und Kapitel 5.2.2.1.3

chen Wachstum von 30%, über das Internet in die Cloud-Ebene übertragen werden, wird für die Berechnung der Anforderungsgebühren in diesem speziellen Fall die Anzahl der Dateien (a_t) berücksichtigt. Es wird angenommen, dass eine Datei die Größe von 1 Megabyte hat. Für die Datenmenge von 1 Terabyte im ersten Jahr existieren demzufolge 1 Millionen Dateien, die in der Cloud-Ebene angelegt werden. Weiterhin wird zur Berechnung der Anforderungskosten angenommen, dass die Anzahl der Dateien ebenfalls um 30% pro Jahr zunimmt. Erfolgte die Zuordnung der Daten zur Ebene „Online“ bzw. zur Klasse mit dem Wert „3“, werden die Speicherplatzkosten der Online-Ebene ($SBK_{Online; t}$) mit dem Betriebskostenfaktor für Verlagerungsvorgänge ($BKF_{V; t}$) und der Datenmenge (m_t) multipliziert. Das resultierende Ergebnis wird zusätzlich mit dem Verlagerungswert⁹⁶⁴ ($V(K_{n; t})$) multipliziert.⁹⁶⁵ Auf diese Weise werden die Verlagerungskosten berechnet, wenn die Daten der Online-Ebene zugeordnet wurden. Zusätzlich muss überprüft werden, ob vor der Verlagerung der Daten in die Ebene „Online“ ein Export der Daten aus der Ebene „Cloud“ stattgefunden hat.⁹⁶⁶ Dazu werden die Datenübertragungskosten für den Export der Daten über das Internet ($VBK_{Ex-Cloud; t}$) mit dem Verlagerungswert für den Cloud-Export⁹⁶⁷ ($V_{CE}(K_{n; t})$) und der Datenmenge (m_t) multipliziert. Die Berechnung der Verlagerungskosten für die Ebenen „Nearline“ und „Offline“ findet auf die gleiche Weise statt, unter Verwendung der Speicherplatzkosten der Nearline-Ebene ($SBK_{Nearline; t}$) und Offline-Ebene ($SBK_{Offline; t}$).⁹⁶⁸

⁹⁶⁴ Vgl. Kapitel 5.2.3.2

⁹⁶⁵ Vgl. hierzu die Vorgehensweise zur Berechnung der Verlagerungskosten Kapitel 5.2.3.3.2.1.1.3

⁹⁶⁶ Amazon berechnet zwar keine Datenübertragungsgebühren, wenn die Daten über das Internet importiert werden. Allerdings werden Gebühren berechnet, wenn die Daten über das Internet aus Amazon S3 exportiert werden. Vgl. hierzu Amazon S3 / S3-Preise / .

⁹⁶⁷ Vgl. Kapitel 5.2.3.2

⁹⁶⁸ Die jeweiligen Speicherplatzkosten werden ebenfalls mit dem Betriebskostenfaktor für Verlagerungsvorgänge, der Datenmenge und dem Verlagerungswert multipliziert.

Die Funktion $VBK_3(K_n)$ berechnet die Betriebskosten für die Verlagerung⁹⁶⁹ der Daten für die betrachteten Datenmengen, die im ersten Jahr 10 Terabyte und 100 Terabyte betragen:⁹⁷⁰

$$\begin{aligned} VBK_3(K_n) = & (WENN(K_{2n;t} = 4; (VBK_{\text{Versand-Cloud};t} \cdot V_{CI}(K_{n;t})); \\ & (WENN(K_{2n;t} = 3; ((SBK_{\text{Online};t} \cdot BKF_{V;t} \cdot m_t) \cdot V(K_{n;t})) + \\ & (VBK_{\text{Versand-Cloud};t} \cdot V_{CE}(K_{n;t})); \\ & (WENN(K_{2n;t} = 2; ((SBK_{\text{Nearline};t} \cdot BKF_{V;t} \cdot m_t) \cdot V(K_{n;t})) + \\ & (VBK_{\text{Versand-Cloud};t} \cdot V_{CE}(K_{n;t})); \\ & ((SBK_{\text{Offline};t} \cdot BKF_{V;t} \cdot m_t) \cdot V(K_{n;t})) + (VBK_{\text{Versand-Cloud};t} \cdot \\ & V_{CE}(K_{n;t})))))) \end{aligned}$$

Es wird angenommen, dass die Daten nicht über das Internet übertragen werden, sondern mittels Versand der Speichermedien über den Postweg. Im Vergleich zu einem Transfer der Daten über das Internet stellt diese Alternative bei großen Datenmengen, wie bspw. 10 oder 100 Terabyte, die kostengünstigere Variante dar.⁹⁷¹ Aufgrund geringer Upload-Geschwindigkeiten von Internetzugängen ist beim Transfer großer Datenmengen mit einem hohen Zeit- und Kostenaufwand zu rechnen. Diese Problematik wird im Rahmen der Wirtschaftlichkeitsanalyse und den durchgeführten Simulationsläufen berücksichtigt. Amazon bietet den direkten Transfer der Daten von einem Speichermedium in das Cloud-System an.⁹⁷² Nachdem das Speichermedium an ein entsprechendes Rechenzentrum versendet wurde, übernimmt Amazon den Import der Daten. Die Kosten⁹⁷³ für den Import der Daten ($VBK_{\text{Versand-Cloud};t}$) werden mit dem Verlagerungswert für den Cloud-Import ($V_{CI}(K_{n;t})$) multipliziert.⁹⁷⁴ Erfolgte die Zuordnung der Daten zur

⁹⁶⁹ Vgl. Kapitel 5.2.3.3.1

⁹⁷⁰ Die Datenmenge beträgt im ersten Jahr 10 bzw. 100 Terabyte und das Datenwachstum beträgt 30% je Folgejahr. Vgl. hierzu Kapitel 5.2.2.1, Anhang 1.

⁹⁷¹ Vgl. Yuan, Yang, Chen / Storage in the Cloud / 20

⁹⁷² Vgl. Lehner, Sattler / Web-Scale Cloud / 174; Amazon / AWS Import und Export /

⁹⁷³ In den Kosten für den Import der Daten durch Amazon sind die Kosten für die Übertragung der Daten in die Cloud-Ebene sowie die Kosten für den Versand, zum Amazon-Rechenzentrum und wieder zurück an den Absender, enthalten. Vgl. hierzu Amazon / Calculator / . Die Administrationskosten, die bspw. für die Vorbereitung der Speichermedien für den Versand entstehen, sind im Betriebskostenfaktor für die Cloud-Ebene (BKF_c) enthalten. Vgl. hierzu Kapitel 5.2.3.3.2.1.2.2, Anhang 5.3.

⁹⁷⁴ Gebühren für die Erstellung von Dateien in Amazon S3 werden in diesem Fall nicht gesondert durch Amazon berechnet. Es fallen nur die Gebühren für den Import der Daten an. Vgl. hierzu Amazon /

Ebene „Online“ bzw. zur Klasse mit dem Wert „3“, werden die Speicherplatzkosten der Online-Ebene ($SBK_{\text{Online}; t}$) mit dem Betriebskostenfaktor für Verlagerungsvorgänge ($BKF_{v; t}$) und der Datenmenge (m_t) multipliziert. Das resultierende Produkt wird zusätzlich mit dem Verlagerungswert ($V(K_{n; t})$) multipliziert. Zusätzlich muss überprüft werden, ob vor der Verlagerung der Daten in die Ebene „Online“ ein Export der Daten aus der Ebene „Cloud“ stattgefunden hat. Dazu werden die Kosten für den Export⁹⁷⁵ der Daten ($VBK_{\text{Versand-Cloud}; t}$) mit dem Verlagerungswert für den Cloud-Export⁹⁷⁶ ($V_{CE}(K_{n; t})$) multipliziert. Die Berechnung der Verlagerungskosten für die Ebenen „Nearline“ und „Offline“ findet auf die gleiche Weise statt, unter Verwendung der Speicherplatzkosten der Nearline-Ebene ($SBK_{\text{Nearline}; t}$) und Offline-Ebene ($SBK_{\text{Offline}; t}$).⁹⁷⁷

5.2.3.3 Interpretation und Ergänzung der Messwerte

Die Klassifizierungsergebnisse bzw. erhobenen Messwerte müssen interpretiert und ergänzt werden, da neben dem Nutzungsgrad der Informationen weitere Kriterien untersucht werden müssen, bevor die Verlagerung⁹⁷⁸ der Informationen auf die jeweils sinnvollste Speicherebene erfolgen kann. Neben rechtlichen Anforderungen oder Kenntnissen über die zukünftige Nutzung von Datenbeständen steht insbesondere die Wirtschaftlichkeit im Vordergrund der Betrachtungen. Die Verlagerung der Informationen ist in Abhängigkeit von der verwendeten Speicherebene mit hohen Zeit- und Kostenaufwänden verbunden.⁹⁷⁹ Die Vorstellung und Interpretation der Klassifizierungsergebnisse der automatisierten Klassifizierung, für die zwei betrachteten Szenarien, sowie die wirtschaftlich sinnvolle Ergänzung bzw. Anpassung dieser Ergebnisse, erfolgen in Kapitel 5.3. Es wird erörtert, für welche Lebenszyklen und Nutzungsgrade die automatisierte Klassifizierung die hinsichtlich der Wirtschaftlichkeit sinnvollsten Ergebnisse liefert. Weiterhin wird untersucht, inwieweit Variationen der Nutzungsgrade, dargestellt durch

Import-Export Preise / .

⁹⁷⁵ Neben dem Import der Daten kann die Option zum Versand der Speichermedien auch für den Export der Daten aus der Cloud-Ebene genutzt werden. Vgl. hierzu Lehner, Sattler / Web-Scale Cloud / 174.

⁹⁷⁶ Vgl. Kapitel 5.2.3.2

⁹⁷⁷ Die jeweiligen Speicherplatzkosten werden ebenfalls mit dem Betriebskostenfaktor für Verlagerungsvorgänge, der Datenmenge und dem Verlagerungswert multipliziert.

⁹⁷⁸ Vgl. Kapitel 5.1.2.4

⁹⁷⁹ Vgl. Kapitel 3.4.1; Bei stark schwankenden Nutzungsgraden können bspw. häufige Verlagerungsvorgänge der Informationen zwischen den Speicherebenen notwendig werden. Vgl. hierzu Kapitel 5.2.3.2.

die unterschiedlich starken Rauschgrade innerhalb der Lebenszyklen,⁹⁸⁰ Auswirkungen auf die Klassifizierungsergebnisse und das Kosten- und Nutzenverhältnis haben.⁹⁸¹

5.3 Demonstration der Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM

Die Demonstration ist dazu geeignet, die Funktionsweise der implementierten Funktionen des Verfahrens zur Wirtschaftlichkeitsanalyse⁹⁸² zu überprüfen bzw. nachzuweisen.⁹⁸³ Die Demonstration ist die Vorstufe der Evaluation und kann als ein Teil der Evaluation betrachtet werden.⁹⁸⁴ Aufbau und Ablauf der Wirtschaftlichkeitsanalyse⁹⁸⁵ wurden bereits in Kapitel 5.2.2 vorgestellt.⁹⁸⁶ In diesem Kapitel wird untersucht, inwieweit durch die Anwendung von ILM Kosten entstehen und welcher Nutzen daraus generiert werden kann.⁹⁸⁷ Hierzu wird eine Kostenvergleichsrechnung⁹⁸⁸ durchgeführt, um die Kosten der Verwaltung⁹⁸⁹ unstrukturierter Daten ohne ILM und der Verwaltung der Daten mit ILM zu vergleichen. Im Rahmen der Kosten- und Nutzenbetrachtungen wird untersucht, welcher monetär messbare Nutzen durch die Anwendung von ILM erzielt werden kann. Wirtschaftlichkeit⁹⁹⁰ von ILM ist gegeben, wenn die Verwaltung der Da-

⁹⁸⁰ Vgl. Kapitel 5.2.2.1.3.1

⁹⁸¹ Hierzu wird die Anzahl der Verlagerungs- bzw. Migrationsvorgänge ermittelt, die gemäß den Ergebnissen der automatisierten Klassifizierung vorgenommen werden müssten. Vgl. hierzu Kapitel 5.2.3.2.

⁹⁸² Vgl. Kapitel 5.2.3

⁹⁸³ Vgl. hierzu Peffers et al. / Research Methodology / 55; Walls, Widmeyer, El Sawy / Design Theory / 36-59.

⁹⁸⁴ Vgl. Venable, Pries-Heje, Baskerville / Framework for Evaluation / 428

⁹⁸⁵ Zur Ausführung und Demonstration der Wirtschaftlichkeitsanalyse wird die Simulationstechnik angewendet. Die Klassifizierung der Daten, die Berechnung der Kosten sowie die Analyse der Wirtschaftlichkeit wird entsprechend simuliert. Vgl. hierzu Kapitel 5.2; Venable, Pries-Heje, Baskerville / Framework for Evaluation / 427; Peffers et al. / Research Methodology / 55. Somit kann für die betrachteten Lebenszyklen, Datenmengen, Kosten und Szenarien die Wirtschaftlichkeitsanalyse im ILM erfolgen. Vgl. Kapitel 5.4 zur Diskussion der verwendeten Methodik für die Demonstration und Evaluation.

⁹⁸⁶ Es wurde erörtert, wie die Funktionen zur Klassifizierung der Daten auf Basis der Nutzungsgrade Klassifizierungswerte berechnen. Die Klassifizierungswerte werden an die Verlagerungs- und Kostenfunktionen übergeben, um die Verlagerungs- und Kostenwerte zu berechnen. Mit Hilfe einer abschließenden Kostenvergleichsrechnung wird die Wirtschaftlichkeit berechnet. Vgl. hierzu Abbildung 5-8, Kapitel 5.2.2. In den Tabellen 5-3 bis 5-14 sind die Ergebnisse der durchgeführten Analyse der Wirtschaftlichkeit abgebildet, diese werden vom implementierten Prototyp in Microsoft Excel ausgegeben.

⁹⁸⁷ Vgl. zur Kosten-Nutzen-Analyse Heinen / Führungslehre / 313-314

⁹⁸⁸ Vgl. Kapitel 4.2.1.1

⁹⁸⁹ Vgl. Kapitel 1.1

⁹⁹⁰ Vgl. Kapitel 4.1

ten im ILM zu Kostensenkungen führt.⁹⁹¹ In Tabelle 5-2 sind die Kosten für die Verwaltung der Daten ohne ILM abgebildet.

Datenmenge [in Terabyte]	Anschaffungs- kosten [in T €]	Speicher- platz- kosten [in T €]	Administ- rations- kosten [in T €]	Verlage- rungs- kosten [in T €]	Gesamt- kosten [in T €]	Verlage- rungen [Anzahl]
1 - 2,856	0	6	58	0	64	0
10 - 28,561	0	60	585	0	644	0
100 - 285,61	0	596	5.845	0	6.441	0

Tab. 5-2: Verwaltung Daten ohne ILM, 5 Jahre

Die Gesamtkosten⁹⁹² ergeben sich für die betrachteten Daten⁹⁹³ aus der Summe von Anschaffungskosten, Speicherplatzkosten, Administrationskosten und Verlagerungskosten.⁹⁹⁴ Die Kosten für Speicherplatz und Administration resultieren aus der Verwaltung der Daten mit Hilfe von Speichermedien mit kurzen Zugriffszeiten.⁹⁹⁵ Die Daten werden bei der Verwaltung ohne ILM nicht verlagert,⁹⁹⁶ eine Betrachtung der Nutzungsgrade bzw. Lebenszyklen findet nicht statt. Es fallen keine Anschaffungskosten für die Entwicklung und Implementierung eines ILM-Konzepts an.⁹⁹⁷ Die in Tabelle 5-2 dargestellten Kosten werden im Folgenden mit den Kosten zur Verwaltung der Daten im ILM

⁹⁹¹ Vgl. Ott / Wirtschaftlichkeitsanalyse / 523; Wiczorrek, Mertens / Management IT-Projekte / 225; Vgl. Kapitel 4.1 zur Definition der Begriffe Kosten und Nutzen. Monetär messbarer Nutzen wird im ILM durch Kostensenkungen erzielt. Vgl. Antweiler / Wirtschaftlichkeitsanalyse / 98ff.; Bechmann / Nutzwertanalyse / 48.

⁹⁹² Vgl. hierzu Tabelle 5-2, „Gesamtkosten“ und Anhang 5.3, Tabelle A5-11 und Tabelle A5-12 zur Berechnung der Speicher- und Administrationskosten zur Verwaltung der Daten ohne ILM. Die Gesamtkosten für die betrachtete Datenmenge 10TB bis 28,56TB (bzw. 100TB bis 285,6TB) betragen das Zehnfache der Gesamtkosten für die Datenmenge 1TB bis 2,856TB (bzw. 10TB bis 28,56TB).

⁹⁹³ Vgl. Kapitel 5.2.2.1

⁹⁹⁴ Vgl. Kapitel 5.2.3.3.1, Tabelle 5-1; Die jeweiligen Kosten wurden für den betrachteten Zeitraum von 5 Jahren bzw. 60 Monaten aufsummiert. Aufgrund des Datenwachstums, den jährlich sinkenden Speicherkosten (Vgl. hierzu Kapitel 5.2.2.1) und den jährlich steigenden Betriebskosten (Vgl. hierzu Anhang 5.3, Tabelle A5-11 und Tabelle A5-12) sind die Administrationskosten, bezogen auf den Zeitraum von 5 Jahren, insgesamt um das 9,8-Fache höher als die Speicherplatzkosten. Vgl. hierzu Glazer / Measuring Information / 99-110; Kaplan, Roy, Srinivasaraghavan / Demand data storage / 1; Forrest, Kaplan, Kindler / Data centers / 1-10; Forrest, Kaplan, Kindler / Energy Efficiency / 1-13; Gantz, Reinsel / Extracting Value / 4; Schadler / Email in the Cloud / 4.

⁹⁹⁵ Vgl. hierzu Tabelle 5-2, „Speicherplatzkosten“ und „Administrationskosten“. Dies entspricht der Nutzung der Speichermedien der Online-Ebene im ILM. Vgl. hierzu Kapitel 5.2.3.1.1.1. Vgl. Anhang 5.3, Tabelle A5-11 und Tabelle A5-12 zur Berechnung der Speicher- und Administrationskosten.

⁹⁹⁶ Die Anzahl der Verlagerungen beträgt für alle Datenmengen null. Vgl. hierzu Tabelle 5-2, „Verlagerungen“. Gleiches gilt für die Verlagerungskosten. Vgl. Tabelle 5-2, „Verlagerungskosten“.

⁹⁹⁷ Vgl. Tabelle 5-2, „Anschaffungskosten“

verglichen.⁹⁹⁸ In den Kapiteln 5.3.1 bis 5.3.6 werden die Kosten für die zwei Szenarien zur Klassifizierung unstrukturierter Daten berechnet. Dabei werden die implementierten Funktionen zur Klassifizierung der Informationen und zur Berechnung der Kosten angewendet. Die Kostenfunktionen werden, bezogen auf die betrachteten Lebenszyklen und Szenarien, für jeden einzelnen Klassifizierungslauf, im Rahmen der Simulation ausgeführt. Bei den Lebenszyklen mit Rauschgraden werden für die jeweiligen Betriebskosten für den Speicherplatz, die Administration und Verlagerung der Daten, aus den Monatswerten der 30 Simulationsläufe, die Mittelwerte berechnet. Die Ergebnisse werden auf den betrachteten Zeitraum von 5 Jahren aufsummiert.

Im ersten Szenario wird die Klasse bzw. Speicherebene „Cloud“⁹⁹⁹ nicht betrachtet. Die Informationen werden im ersten Szenario ausschließlich auf unternehmensinternen Speichermedien verwaltet.¹⁰⁰⁰ Im zweiten Szenario werden neben den unternehmensinternen Speichermedien zusätzlich Public-Cloud-Dienste bzw. IaaS-Dienste genutzt.¹⁰⁰¹ Die in den Tabellen 5-3 bis 5-14 enthaltenen Anschaffungs-, Speicherplatz-, Administrations- und Verlagerungskosten stellen die Gesamtkosten für die jeweiligen Nutzungsgrade¹⁰⁰² dar.¹⁰⁰³ Durch den Vergleich der Kosten beider Szenarien mit den Kosten zur Verwaltung der Daten ohne ILM¹⁰⁰⁴ wird untersucht,¹⁰⁰⁵ inwiefern ILM zu einer Kostensenkung beiträgt und ob die Nutzung von Cloud-Diensten im ILM zusätzliche Auswirkungen auf das Kosten- und Nutzenverhältnis bezüglich der Verwaltung unstrukturierter Daten hat.¹⁰⁰⁶

⁹⁹⁸ Der Vergleich findet für beide ILM-Szenarien statt.

⁹⁹⁹ Vgl. Kapitel 3.4.2

¹⁰⁰⁰ Vgl. Kapitel 5.2.3.1.1.1.1

¹⁰⁰¹ Vgl. Kapitel 5.2.3.1.1.1.2

¹⁰⁰² Vgl. Kapitel 5.2.2.1.3.1

¹⁰⁰³ Vgl. hierzu Tabelle 5-3 bis 5-14, „Anschaffungskosten“, Speicherplatzkosten“, „Administrationskosten“, „Verlagerungskosten“ und „Gesamtkosten“ für die betrachteten sechs Nutzungsgrade bzw. Lebenszyklen. Vgl. Kapitel 5.2.2.1.3.1.

¹⁰⁰⁴ Vgl. Kapitel 5.3, Tabelle 5-2

¹⁰⁰⁵ Für die Kostenvergleichsrechnung und die Berechnung der Kosten gelten die Annahmen zu den Kostenarten bzw. Betriebskosten. Vgl. hierzu Kapitel 5.2.3.3.1 und Anhang 5.3.

¹⁰⁰⁶ Vgl. hierzu Tabelle 5-3 bis 5-14, „Kostenveränderung“. Für die Berechnung der Kostensenkung bzw. Kostenveränderung in Prozent dient der jeweilige Gesamtkostenwert zur Verwaltung der Daten ohne ILM als Grundwert. Vgl. hierzu Kapitel 5.3, Tabelle 5-2, Gesamtkostenwerte „64“, „644“ und „6441“. Die Gesamtkostenwerte der Verwaltung der Daten mit ILM bilden die Prozentwerte (Vgl. Tabelle 5-3 bis 5-14). Vgl. zur Verwendung von Grund- und Prozentwert zur Berechnung von Kostenabweichungen Kem-

5.3.1 Sinkender Nutzungsgrad

Beim sinkenden Nutzungsgrad erfolgen die häufigsten Zugriffe kurz nach dem Entstehen der Information, danach wird in der Regel nur noch selten auf die Information zugegriffen.¹⁰⁰⁷ In Tabelle 5-3 sind die Kosten für den sinkenden Nutzungsgrad enthalten, die bei der automatisierten Verwaltung unstrukturierter Daten im ILM für das erste Szenario entstehen.

Rauschgrad [in T]	Anschaffungs- kosten [in T €]	Speicher- platz- kosten [in T €]	Administra- tions- kosten [in T €]	Verlage- rungs- kosten [in T €]	Verlage- rungen [Anzahl]	Gesamt- kosten [in T €]	Kosten- verände- rung [in %]
Datenmenge 1 - 2,856							
0	60	1,7	14	0,1	3	76	17
5	60	2,9	37	1,9	35	102	58
10	60	3,0	42	2,1	34	107	66
20	60	3,2	45	2,0	32	111	72
Datenmenge 10 - 28,561							
0	60	17	139	0,8	3	217	-66
5	60	29	367	19	35	475	-26
10	60	30	417	21	34	528	-18
20	60	32	455	20	32	567	-12
Datenmenge 100 - 285,61							
0	60	171	1.388	8	3	1.627	-75
5	60	288	3.671	194	35	4.213	-35
10	60	305	4.173	206	34	4.744	-26
20	60	319	4.546	204	32	5.129	-20

Tab. 5-3: Sinkender Nutzungsgrad, ILM ohne Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad

Die Gesamtkosten betragen bei der ersten Datenmenge¹⁰⁰⁸ für den Nutzungsgrad ohne Rauschgrad¹⁰⁰⁹ ca. 76.000 Euro. Verglichen mit den Gesamtkosten zur Verwaltung der Daten ohne ILM¹⁰¹⁰ stellt dies eine Erhöhung der Kosten um 17 Prozent dar.¹⁰¹¹ Ob-

metmüller, Bogensberger / Kostenrechnung / 28-31. Somit wird berechnet, um wieviel Prozent die Kosten aufgrund der Verwaltung der Daten im ILM, bezogen auf die Kosten zur Verwaltung der Daten ohne ILM, gesunken oder gestiegen sind.

¹⁰⁰⁷ Vgl. Kapitel 5.2.2.1.3.1.1

¹⁰⁰⁸ Bei einer Datenmenge von 1TB bis 2,856TB.

¹⁰⁰⁹ Vgl. Tabelle 5-3, Rauschgrad „0“

¹⁰¹⁰ Vgl. Kapitel 5.3, Tabelle 5-2

wohl die Betriebskosten bzw. die Speicherplatz- und Administrationskosten mit ILM gesenkt wurden,¹⁰¹² ist die Verwaltung der Daten im ILM wirtschaftlich nicht sinnvoll. Der Grund hierfür sind die Anschaffungskosten für die Entwicklung und Implementierung des ILM-Konzepts,¹⁰¹³ die die Senkung der Betriebskosten übersteigen. Werden die Rauschgrade erhöht und die Nutzungsgrade unregelmäßiger, so steigen die Betriebskosten und das Kostensenkungspotential von ILM nimmt weiter ab.¹⁰¹⁴ Durch das Verändern der Rauschgrade müssen die Daten der Nutzungsgrade entsprechend auf den Ebenen Online, Nearline und Offline abgelegt werden. Nehmen die Zugriffe durch die Erhöhung der Rauschgrade zu, so werden die Daten in der kostenintensiven Online-Ebene abgelegt.¹⁰¹⁵ Eine durchgängige Speicherung der Daten auf der kostengünstigen Offline-Ebene ab dem zweiten Jahr ist nicht mehr möglich, wie es bei dem Nutzungsgrad ohne Rauschgrad der Fall war. Da die Daten bei zunehmenden Schwankungen der Nutzungsgrade öfter verlagert werden müssen, steigen zusätzlich die Anzahl der Verlagerungsvorgänge und die damit verbundenen Verlagerungskosten. Bei einem Rauschgrad von zwanzigtausend sinkt die Anzahl der Verlagerungen geringfügig, da die Daten aufgrund des hohen Rauschgrads länger bzw. über mehrere Monate hinweg in ein und derselben Speicherebene gehalten werden.

Die Betriebskosten sind im ILM bei der zweiten Datenmenge¹⁰¹⁶ um das Zehnfache höher, als bei der ersten Datenmenge.¹⁰¹⁷ Die Anschaffungskosten bleiben gleich. Die Schulungskosten und die Kosten für die Implementierung der Funktionen zur automatisierten Klassifizierung entstehen unabhängig von der betrachteten Datenmenge.¹⁰¹⁸ Eine

¹⁰¹¹ Vgl. hierzu Tabelle 5-3, Kostenveränderung „17%“. Sind die Kosten mit ILM niedriger als die Kosten zur Verwaltung der Daten ohne ILM, so wird dem Wert der Kostensenkung bzw. Kostenveränderung ein negatives Vorzeichen vorangestellt.

¹⁰¹² Im Vergleich zur Verwaltung der Daten ohne ILM, siehe Tabelle 5-2. Ab dem zweiten Jahr können die Daten von der Online-Ebene zur Offline-Ebene verlagert werden, da keine Zugriffe auf die Daten erfolgen. Vgl. hierzu Kapitel 5.2.2.1.3.1.1. Deshalb können die Speicherplatz- und Administrationskosten gesenkt werden. Die Kosten für die Verlagerung der Daten sind gering, da lediglich drei Verlagerungen stattfinden.

¹⁰¹³ Vgl. Kapitel 5.2.3.3.1.1; Anhang 5.1

¹⁰¹⁴ Vgl. hierzu Tabelle 5-3, Kostenveränderung „58%“, „66%“, „72%“.

¹⁰¹⁵ Vgl. hierzu die Intervalle der Klassifizierungsfunktionen Kapitel 5.2.3.1.1.1.

¹⁰¹⁶ Bei einer Datenmenge von 10TB bis 28,56TB.

¹⁰¹⁷ Vgl. hierzu Tabelle 5-3, „Speicherplatzkosten“, „Administrationskosten“ und „Verlagerungskosten“. Die Anzahl der Verlagerungen bleibt gleich, da diese unabhängig von der Datenmenge ist. Vgl. hierzu Kapitel 5.2.3.2.

¹⁰¹⁸ Vgl. Anhang 5.1

Abhängigkeit von der Datenmenge ist bspw. bei den Kosten für die Vorklassifizierung der Daten und die Zuordnung der Daten zu Dokumentenklassen gegeben. Diese Kosten werden den Betriebskosten zugeordnet.¹⁰¹⁹ Die Kostensenkung beträgt bei einem Rauschgrad von „null“ 66 Prozent.¹⁰²⁰ Die Wirtschaftlichkeit von ILM ist gegeben. Ähnlich wie bei der ersten Datenmenge sinkt das Kostensenkungspotential bei einer Erhöhung der Rauschgrade, die Wirtschaftlichkeit von ILM ist allerdings weiterhin gegeben.¹⁰²¹ Die Betriebskosten der dritten Datenmenge¹⁰²² sind im ILM um das Zehnfache höher als die Betriebskosten der zweiten Datenmenge.¹⁰²³ Die Kostensenkung beträgt bei einem Rauschgrad von „null“ 75 Prozent.¹⁰²⁴ Das Kostensenkungspotential ist bei der dritten Datenmenge höher, als bei der zweiten Datenmenge. Das liegt neben der Senkung der Speicherplatz- und Betriebskosten daran, dass die Anschaffungskosten unabhängig vom Datenwachstum sind und nur die Betriebskosten entsprechend des Datenwachstums steigen. Demzufolge sinkt der Anteil der Anschaffungskosten an den Gesamtkosten bei einer wachsenden Datenmenge. Die Wirtschaftlichkeit sinkt bei der dritten Datenmenge ebenfalls bei Erhöhung der Rauschgrade. In Tabelle 5-4 sind die Kosten für den sinkenden Nutzungsgrad enthalten, die im ILM für das zweite Szenario entstehen.

¹⁰¹⁹ Vgl. hierzu Kapitel 5.2.3.3.1 und Anhang 5.3.

¹⁰²⁰ Verglichen mit den Kosten, die bei der Verwaltung der Datenmenge ohne ILM entstehen würden. Vgl. hierzu Kapitel 5.3, Tabelle 5-2, Gesamtkosten „644“ sowie Tabelle 5-3, Kostenveränderung „-66%“.

¹⁰²¹ Vgl. hierzu Tabelle 5-3, Kostenveränderung „-26%“, „-18%“, „-12%“.

¹⁰²² Bei einer Datenmenge von 100TB bis 285,6TB.

¹⁰²³ Vgl. hierzu Tabelle 5-3, „Speicherplatzkosten“ und „Administrationskosten“. Die Anzahl der Verlagerungen bleibt gleich, da diese unabhängig von der Datenmenge ist. Vgl. hierzu Kapitel 5.2.3.2.

¹⁰²⁴ Verglichen mit den Kosten, die bei der Verwaltung der Datenmenge ohne ILM entstehen würden. Vgl. hierzu Kapitel 5.3, Tabelle 5-2, Gesamtkosten „6.441“ sowie Tabelle 5-3, Gesamtkosten „1.627“ bzw. Kostenveränderung „-75%“.

Rauschgrad [in T]	Anschaffungs- kosten [in T €]	Speicher- platz- kosten [in T €]	Administra- tions- kosten [in T €]	Verlage- rungs- kosten [in T €]	Verlage- rungen [Anzahl]	Gesamt- kosten [in T €]	Kosten- verände- rung [in %]
Datenmenge 1 - 2,856							
0	75	1,2	10	0,2	6	87	35
5	75	2,2	30	2,5	49	109	70
10	75	1,8	23	3,3	61	103	60
20	75	1,3	15	3,4	65	94	46
Datenmenge 10 - 28,561							
0	75	12	105	1,2	6	193	-70
5	75	22	295	20	49	412	-36
10	75	18	230	19	61	342	-47
20	75	13	146	15	65	249	-61
Datenmenge 100 - 285,61							
0	75	123	1.049	10	6	1.257	-80
5	75	225	2.952	191	49	3.443	-47
10	75	181	2.300	176	61	2.732	-58
20	75	135	1.457	129	65	1.796	-72

Tab. 5-4: Sinkender Nutzungsgrad, ILM mit Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad

Bezüglich der ersten Datenmenge ist die Verwaltung der Daten wirtschaftlich nicht sinnvoll.¹⁰²⁵ Die Speicherplatz- und Administrationskosten sind geringer als bei der Verwaltung der Daten im ersten Szenario,¹⁰²⁶ aber die Anschaffungskosten und Verlagerungskosten sind im zweiten Szenario aufgrund der zusätzlichen Verwendung der Cloud-Ebene höher.¹⁰²⁷ Die Wirtschaftlichkeit von ILM ist für die weiteren Datenmengen gegeben. Das Kostensenkungspotential ist bei einem Rauschgrad von „null“ am größten.¹⁰²⁸ Steigt der Rauschgrad von „null“ auf „fünftausend“, so sinkt das Kostensenkungspotential bei beiden Datenmengen.¹⁰²⁹ Allerdings nimmt anschließend das

¹⁰²⁵ Vgl. hierzu Tabelle 5-4, Kostenänderung „35%“, „70%“, „60%“, „46%“.

¹⁰²⁶ Vgl. hierzu Tabelle 5-3 und Tabelle 5-4, „Speicherplatzkosten“ und „Administrationskosten“.

¹⁰²⁷ Vgl. Anhang 5.1; Anhang 5.3

¹⁰²⁸ Neben dem Rauschgrad hat die Datenmenge ebenfalls Einfluss auf die Wirtschaftlichkeit. Je größer die Datenmenge ist, desto höher ist das Kostensenkungspotential. Vgl. hierzu Tabelle 5-4, Kostenänderung „-70%“ für die Datenmenge von 10TB bis 28,56TB und Kostenänderung „-80%“ für die Datenmenge von 100TB bis 285,6TB.

¹⁰²⁹ Vgl. hierzu Tabelle 5-4, Kostenänderung „-70%“ ohne Rauschgrad und „-36%“ mit einem Rauschgrad von „5000“ für die Datenmenge 10TB bis 28,56TB sowie Kostenänderung „-80%“ ohne Rauschgrad und „-47%“ mit einem Rauschgrad von „5000“ für die Datenmenge 100TB bis 285,6TB. Ähnlich wie bei der Verwaltung der Daten des sinkenden Nutzungsgrads im ersten Szenario steigen die Betriebskosten durch die Veränderung der Zugriffe, da ab dem zweiten Jahr nicht mehr ausschließlich die Offline-Ebene

Kostensenkungspotential mit steigendem Rauschgrad wieder zu.¹⁰³⁰ Der Grund hierfür ist, dass bei höheren Nutzungsgraden, hervorgerufen durch höhere Rauschgrade, die Daten öfter der Cloud-Ebene zugewiesen werden.¹⁰³¹ Die Klassifizierung und Verlagerung der Daten erfolgt auf Monatsbasis. Für den Zeitraum von 5 Jahren werden die Klassifizierungs- und Verlagerungsfunktionen 60 Mal aufgerufen. Dennoch ist es möglich, dass in den 5 Jahren mehr als 60 Verlagerungsvorgänge stattfinden. Wenn bspw. die Daten von der Cloud-Ebene auf ein unternehmensinternes Speichermedium verlagert werden sollen, müssen die Daten zunächst aus der Cloud-Ebene exportiert werden, bspw. über den Versand auf dem Postweg. Da in diesem Fall keine direkte Datenübertragung zwischen den Speicherebenen möglich ist, so wie es bei der Verlagerung von Daten zwischen unternehmensinternen Speichermedien der Fall wäre, wird der Datenexport aus der Cloud gesondert als Verlagerungsvorgang gezählt.¹⁰³² Durch die Nutzung von Public-Cloud-Diensten bzw. Cloud-Speicher (IaaS) können Lastspitzen hinsichtlich des Nutzungsgrads abgedeckt werden. Die Speicherplatz-, Administrations-, und Verlagerungskosten der Cloud-Ebene sind niedriger als die jeweiligen Kosten der anderen Speicherebenen.¹⁰³³ Trotz der steigenden Anzahl der Verlagerungsvorgänge, sinken die Kosten für die Verlagerungen. Der Versand der Daten großer Datenmengen über den Postweg ist günstiger, als der Transfer der Daten zwischen den unternehmenseigenen Speichermedien.¹⁰³⁴ Die Verwaltung der Daten des sinkenden Nutzungsgrads im zweiten Szenario führt zu einer höheren Kostensenkung als die Verwaltung der Daten im ersten Szenario.¹⁰³⁵

genutzt wird und die Kosten für die Verlagerungsvorgänge steigen.

¹⁰³⁰ Vgl. hierzu Tabelle 5-4, Kostenänderung „-36%“, „-47%“, „-61%“ sowie „-47%“, „-58%“, „-72%“ für die jeweiligen Datenmengen und Rauschgrade.

¹⁰³¹ Vgl. hierzu die Intervalle der Klassifizierungsfunktionen Kapitel 5.2.3.1.1.1.

¹⁰³² Vgl. hierzu Kapitel 5.2.3.2 und 5.2.3.3.2.1.2.

¹⁰³³ Vgl. hierzu Anhang 5.3; Iannucci, Gupta / Cloud Data Center / 15; Pallickara, Pallickara, Pierce / Data Management in the Cloud / 528; Mahmood, Saeed / Frameworks for the Cloud / 314-315; Jewell et al. / Performance and Capacity / 10-11.

¹⁰³⁴ Vgl. Lehner, Sattler / Web-Scale Cloud / 174; Chee, Franklin / Cloud Strategies / 119; Anhang 5.3

¹⁰³⁵ Vgl. hierzu Kostenänderung für die Datenmengen und Rauschgrade Tabelle 5-3 und Tabelle 5-4.

5.3.2 Steigender Nutzungsgrad

Beim steigenden Nutzungsgrad sinken die Zugriffe kurz nach dem Entstehen der Daten und ab dem zweiten Jahr nehmen die Zugriffe wieder zu.¹⁰³⁶ In Tabelle 5-5 sind die Kosten für den steigenden Nutzungsgrad enthalten, die bei der automatisierten Verwaltung der Daten im ILM für das erste Szenario entstehen.

Rauschgrad [in T]	Anschaffungs- kosten [in T €]	Speicher- platz- kosten [in T €]	Administra- tions- kosten [in T €]	Verlage- rungs- kosten [in T €]	Verlage- rungen [Anzahl]	Gesamt- kosten [in T €]	Kosten- verände- rung [in %]
Datenmenge 1 - 2,856							
0	60	4,2	75	0,2	5	139	117
5	60	4,1	71	1,2	24	136	112
10	60	3,8	63	1,8	30	128	99
20	60	3,4	54	1,9	31	120	86
Datenmenge 10 - 28,561							
0	60	42	751	2	5	855	33
5	60	41	710	12	24	824	28
10	60	38	628	18	30	743	15
20	60	34	544	19	31	657	2
Datenmenge 100 - 285,61							
0	60	419	7.509	19	5	8.006	24
5	60	413	7.101	122	24	7.696	19
10	60	379	6.277	177	30	6.893	7
20	60	341	5.436	190	31	6.027	-6

Tab. 5-5: Steigender Nutzungsgrad, ILM ohne Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad

Die Verwaltung der Daten des steigenden Lebenszyklus im ILM ist mit höheren Kosten verbunden, als die Verwaltung der Daten ohne ILM.¹⁰³⁷ So wie beim sinkenden Nutzungsgrad gilt für den steigenden Nutzungsgrad dennoch, dass das Kostensenkungspotential mit wachsender Datenmenge und steigendem Rauschgrad zunimmt. Die Daten werden durch die Klassifizierung im ILM aufgrund des steigenden Nutzungsgrads vorwiegend der Online-Ebene zugeordnet.¹⁰³⁸ Das Kostensenkungspotential von ILM geht

¹⁰³⁶ Vgl. Kapitel 5.2.2.1.3.1.2

¹⁰³⁷ Vgl. hierzu Tabelle 5-5, Kostenänderung für die betrachteten Datenmengen und Rauschgrade.

¹⁰³⁸ Ab einer Anzahl von 3000 Zugriffen erfolgt die Zuordnung zur Online-Ebene. Vgl. hierzu die Klassifizierungsfunktion für Szenario 1, Kapitel 5.2.3.1.1.1.1. Die Daten werden ab der zweiten Jahreshälfte des zweiten Jahres der Online-Ebene zugeordnet.

verloren,¹⁰³⁹ da die kostengünstigeren Speicherebenen „Nearline“ und „Offline“ weniger zum Einsatz kommen. Die Administrationskosten sind im ILM ebenfalls höher, so dass die Verwaltung der Daten auf der Online-Ebene im ILM mehr Kosten verursacht, als die Verwaltung der Daten auf der Online-Ebene ohne den Einsatz von ILM.¹⁰⁴⁰ Mit dem Anstieg der Rauschgrade sinken die Speicherplatz- und Administrationskosten, die Daten werden nun häufiger der Nearline-Ebene oder Offline-Ebene zugeordnet. Dennoch ist die Wirtschaftlichkeit nicht gegeben. Eine Ausnahme ist die dritte Datenmenge¹⁰⁴¹ bei einem Rauschgrad von zwanzigtausend. In diesem Fall ist die Wirtschaftlichkeit von ILM gegeben. Durch den starken Rauschgrad wird die Grundfunktion des steigenden Rauschgrads derart verändert, dass die Zugriffe bzw. Nutzungsgrade zusätzlich den Intervallen der Speicherebenen „Nearline“ und „Offline“ zugeordnet werden können. Die Kostenänderung bzw. Kostensenkung beträgt, im Vergleich zur Verwaltung der Daten ohne ILM, sechs Prozent. Allerdings ist zu beachten, dass der steigende Nutzungsgrad aufgrund des hohen Rauschgrads entsprechend stark von der ursprünglichen Grundfunktion abweicht.¹⁰⁴² Unter Berücksichtigung der getroffenen Annahmen zu den Klassifizierungsfunktionen¹⁰⁴³ und Betriebskosten¹⁰⁴⁴ ist davon auszugehen, dass ILM ohne Cloud-Ebene für den steigenden Nutzungsgrad nur wirtschaftlich sinnvoll ist, wenn große Datenmengen verwaltet werden und wenn der steigende Nutzungsgrad sehr hohe Schwankungen aufweist. In Tabelle 5-6 sind die Kosten für den steigenden Nutzungsgrad enthalten, die im ILM für das zweite Szenario entstehen.

¹⁰³⁹ Vgl. hierzu Tabelle 5-5, Kostenänderung für die betrachteten Datenmengen und Rauschgrade.

¹⁰⁴⁰ Vgl. Anhang 5.3, Tabelle A5-4 und Tabelle A5-12

¹⁰⁴¹ Vgl. Tabelle 5-5, Datenmenge 100TB bis 285,6TB

¹⁰⁴² Vgl. hierzu den steigenden Nutzungsgrad mit entsprechendem Rauschgrad von „20000“ in Abbildung 5-10.

¹⁰⁴³ Vgl. Kapitel 5.2.3.1

¹⁰⁴⁴ Vgl. Kapitel 5.2.3.3.1, Anhang 5.3

Rauschgrad [in T]	Anschaffungs- kosten [in T €]	Speicher- platz- kosten [in T €]	Administra- tions- kosten [in T €]	Verlage- rungs- kosten [in T €]	Verlage- rungen [Anzahl]	Gesamt- kosten [in T €]	Kosten- verände- rung [in %]
Datenmenge 1 - 2,856							
0	75	2,7	30	0,2	7	108	67
5	75	2,4	26	3,0	52	107	66
10	75	1,8	20	3,5	64	101	56
20	75	1,4	15	3,3	64	94	46
Datenmenge 10 - 28,561							
0	75	27	296	2	7	400	-38
5	75	24	265	18	52	381	-41
10	75	18	201	18	64	312	-52
20	75	14	146	14	64	249	-61
Datenmenge 100 - 285,61							
0	75	269	2.962	21	7	3.326	-48
5	75	235	2.646	167	52	3.123	-52
10	75	185	2.012	162	64	2.433	-62
20	75	138	1.457	127	64	1.797	-72

Tab. 5-6: Steigender Nutzungsgrad, ILM mit Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad

ILM führt hinsichtlich der ersten Datenmenge zu einer Erhöhung der Kosten.¹⁰⁴⁵ Wirtschaftlichkeit von ILM ist für die weiteren Datenmengen gegeben. Mit zunehmendem Wachstum der Daten und steigendem Rauschgrad erhöht sich das Kostensenkungspotential.¹⁰⁴⁶ Die Daten werden aufgrund des steigenden Nutzungsgrads im zweiten Jahr auf die Online-Ebene verlagert.¹⁰⁴⁷ Wird eine Anzahl von achttausend Zugriffen erreicht, werden die Daten der Cloud-Ebene zugeordnet.¹⁰⁴⁸ Durch die Erhöhung der Rauschgrade steigt die Anzahl der Verlagerungsvorgänge, die Kosten für die Verlagerungen sinken. Die Verlagerungskosten steigen bei einem Rauschgrad von fünftausend zunächst deutlich an, da die Daten vorwiegend der Online-Ebene zugeordnet werden. Ab einem Rauschgrad von zehntausend sinken die Verlagerungskosten, da die Daten über den kostengünstigeren Cloud-Import und Cloud-Export verlagert werden. Die Anzahl der Verlagerungen ist für die Rauschgrade von zehntausend und zwanzigtausend

¹⁰⁴⁵ Vgl. hierzu Tabelle 5-6, Kostenänderung „67%“, „66%“, „56%“, „46%“.

¹⁰⁴⁶ Vgl. Tabelle 5-6, Kostenänderung der drei Datenmengen

¹⁰⁴⁷ Führen die Rauschgrade zu einer Verringerung der Anzahl der Zugriffe, führt dies dazu, dass die Daten der Nearline-Ebene oder Offline-Ebene zugeordnet werden. Das wirkt sich positiv auf die Wirtschaftlichkeit im ILM aus.

¹⁰⁴⁸ Vgl. zu den Intervallen der Speicherebenen und Klassifizierungsfunktionen Kapitel 5.2.3.1.1.1.2.

gleich.¹⁰⁴⁹ Bei einem Rauschgrad von zehntausend nimmt die Anzahl der Verlagerungen, im Vergleich zum Rauschgrad von fünftausend, zu. Für den Rauschgrad von zwanzigtausend gilt, dass in den Jahren 1 und 2 die Anzahl der Verlagerungen weiter steigt, aber in den Jahren 4 und 5 die Daten für längere Zeiträume in der Cloud-Ebene gehalten werden. Die Speicherplatz- und Administrationskosten sinken mit steigendem Rauschgrad.¹⁰⁵⁰ Die Verwaltung der Daten des steigenden Nutzungsgrads im zweiten Szenario führt zu einer höheren Kostensenkung als die Verwaltung der Daten im ersten Szenario.

5.3.3 Konstanter Nutzungsgrad

Der konstante Nutzungsgrad beschreibt einen Lebenszyklus, bei dem die Anzahl der Zugriffe, über den betrachteten Zeitraum von fünf Jahren, konstant bleibt. Tabelle 5-7 enthält die Kosten für die Verwaltung der Daten des konstanten Nutzungsgrads für das erste Szenario. Die Daten der Nutzungsgrade ohne Rauschgrad werden über den gesamten Zeitraum von fünf Jahren der Online-Ebene zugeordnet.¹⁰⁵¹ Eine Verlagerung der Daten findet nicht statt.¹⁰⁵² Mit zunehmendem Rauschgrad und wachsender Datenmenge werden die Daten zusätzlich den Speicherebenen „Nearline“ und „Offline“ zugeordnet. Die Kosten sinken, ILM ist dennoch nicht wirtschaftlich.

¹⁰⁴⁹ Vgl. hierzu Tabelle 5-6, Verlagerungen Anzahl „64“.

¹⁰⁵⁰ Da durch eine hohe Anzahl an Zugriffen öfter die Cloud-Ebene verwendet wird. Vgl. hierzu Tabelle 5-6, Speicherplatzkosten, Administrationskosten.

¹⁰⁵¹ Vgl. hierzu die Beschreibung der Grundfunktion des konstanten Nutzungsgrads, Kapitel 5.2.2.1.3.1.3. Die Anzahl der Zugriffe beträgt konstant fünftausend. Vgl. zur Klassifizierung der Daten und Zuordnung zur Online-Ebene mit Hilfe der Klassifizierungsfunktionen Kapitel 5.2.3.1.1.1.1.

¹⁰⁵² Bezüglich der drei Datenmengen und einem Rauschgrad von „null“ wird in Tabelle 5-7 jeweils ein Verlagerungsvorgang angegeben. Es gilt die Annahme, dass die Daten vor der Implementierung des ILM-Konzepts und der Durchführung der automatisierten Klassifizierung auf anderen Speichermedien gespeichert waren und im jeweiligen ersten Monat zunächst auf die Online-Ebene verlagert werden müssen. Vgl. hierzu die implementierten Verlagerungsfunktionen Kapitel 5.2.3.2.

Rauschgrad [in T]	Anschaffungs- kosten [in T €]	Speicher- platz- kosten [in T €]	Administra- tions- kosten [in T €]	Verlage- rungs- kosten [in T €]	Verlage- rungen [Anzahl]	Gesamt- kosten [in T €]	Kosten- verände- rung [in %]
Datenmenge 1 - 2,856							
0	60	6,0	88	0,1	1	154	139
5	60	4,3	63	2,3	33	129	101
10	60	3,7	55	2,2	35	121	88
20	60	3,6	53	2,1	33	119	84
Datenmenge 10 - 28,561							
0	60	60	877	1	1	997	55
5	60	43	627	23	33	752	17
10	60	37	554	22	35	672	4
20	60	36	531	21	33	648	1
Datenmenge 100 - 285,61							
0	60	596	8.768	6	1	9.430	46
5	60	428	6.267	228	33	6.983	8
10	60	373	5.535	216	35	6.184	-4
20	60	357	5.310	209	33	5.936	-8

Tab. 5-7: Konstanter Nutzungsgrad, ILM ohne Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad

Wirtschaftlichkeit ist, ähnlich wie beim steigenden Nutzungsgrad, für die dritte Datenmenge erst gegeben, wenn der Rauschgrad entsprechend hoch ist.¹⁰⁵³ Diese Funktionen weichen aufgrund der Rauschgrade entsprechend stark von der ursprünglichen Grundfunktion ab, sodass ein konstanter Trend nicht mehr vorliegt. Unter Berücksichtigung der getroffenen Annahmen zu den Klassifizierungsfunktionen und Betriebskosten ist deshalb davon auszugehen, dass ILM ohne Cloud-Ebene für den konstanten Nutzungsgrad nicht wirtschaftlich ist. Entscheidend dafür ist der entsprechende Nutzungsgrad bzw. die Anzahl der Zugriffe und die Zuordnung der Daten zu den Speicherebenen. Aufgrund der konstanten Zugriffe von fünftausend Zugriffen ordnen die Klassifizierungsfunktionen die Daten der Online-Ebene zu. Bei einer geringeren Anzahl von konstanten Zugriffen, wie bspw. eintausend oder zweitausend Zugriffen, würden die Daten der kostengünstigeren Nearline-Ebene zugeordnet werden. In diesem Fall würde ILM zu Kostensenkungen führen, da im Gegensatz zur Verwaltung der Daten ohne ILM kostengünstigere Speichermedien zum Einsatz kämen.

¹⁰⁵³ Vgl. hierzu Tabelle 5-7, Datenmenge 100TB bis 285,6TB, Rauschgrade „10000“ und „20000“.

Rauschgrad [in T]	Anschaf- fungs- kosten [in T €]	Speicher- platz- kosten [in T €]	Administra- tions- kosten [in T €]	Verlage- rungs- kosten [in T €]	Verlage- rungen [Anzahl]	Gesamt- kosten [in T €]	Kosten- verände- rung [in %]
Datenmenge 1 - 2,856							
0	75	6,0	88	0,1	1	169	162
5	75	3,0	40	4,4	69	122	90
10	75	1,9	24	4,0	71	105	63
20	75	1,4	15	3,6	64	95	48
Datenmenge 10 - 28,561							
0	75	60	877	1	1	1.012	57
5	75	30	397	27	69	529	-18
10	75	19	238	20	71	353	-45
20	75	14	153	16	64	258	-60
Datenmenge 100 - 285,61							
0	75	596	8.768	6	1	9.445	47
5	75	298	3.969	254	69	4.597	-29
10	75	192	2.383	187	71	2.837	-56
20	75	141	1.534	138	64	1.887	-71

Tab. 5-8: Konstanter Nutzungsgrad, ILM mit Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad

In Tabelle 5-8 sind die Kosten für den konstanten Nutzungsgrad enthalten, die im ILM für das zweite Szenario entstehen. Die Daten der Nutzungsgrade ohne Rauschgrad werden über den gesamten Zeitraum von fünf Jahren der Online-Ebene zugeordnet.¹⁰⁵⁴ Für die jährlich wachsenden Datenmengen von 10 Terabyte und 100 Terabyte ist das ILM mit Cloud-Ebene, für die Nutzungsgrade mit Rauschgrad, wirtschaftlich.¹⁰⁵⁵ Trotz zusätzlicher Verwendung der Cloud-Ebene werden bei der ersten Datenmenge keine Kostensenkungen erzielt. Die Verwaltung der Daten im zweiten Szenario ist im Gegensatz zum ersten Szenario, für die Datenmengen zwei und drei sowie den Nutzungsgraden mit Rauschgrad, wirtschaftlich.

¹⁰⁵⁴ Vgl. hierzu die Erläuterungen zu Tabelle 5-7. Aus diesem Grund ist ILM für den Nutzungsgrad ohne Rauschgrad bzw. für die Grundfunktion des konstanten Nutzungsgrads nicht wirtschaftlich.

¹⁰⁵⁵ Je höher der Rauschgrad ist, desto höher ist die Kostensenkung.

5.3.4 Saisonaler Nutzungsgrad

Die Kosten für die Verwaltung der Daten des saisonalen Nutzungsgrads ohne Cloud-Ebene sind in Tabelle 5-9 dargestellt.

Rauschgrad [in T]	Anschaffungs- kosten [in T €]	Speicher- platz- kosten [in T €]	Administra- tions- kosten [in T €]	Verlage- rungs- kosten [in T €]	Verlage- rungen [Anzahl]	Gesamt- kosten [in T €]	Kosten- verände- rung [in %]
Datenmenge 1 - 2,856							
0	60	4,4	64	1,1	21	130	101
5	60	4,3	64	1,3	22	130	101
10	60	4,0	59	1,8	29	125	94
20	60	3,7	55	2,0	32	120	87
Datenmenge 10 - 28,561							
0	60	44	643	11	21	757	18
5	60	43	638	13	22	755	17
10	60	40	594	18	29	712	11
20	60	37	547	20	32	664	3
Datenmenge 100 - 285,61							
0	60	437	6.430	106	21	7.033	9
5	60	435	6.382	134	22	7.011	9
10	60	400	5.938	181	29	6.579	2
20	60	370	5.469	203	32	6.101	-5

Tab. 5-9: Saisonaler Nutzungsgrad, ILM ohne Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad

Aufgrund der saisonalen Schwankungen der Zugriffe werden die Daten den Speicher-ebenen „Online“, „Nearline“ und „Offline“ zugeordnet.¹⁰⁵⁶ Die Speicherplatzkosten werden gesenkt.¹⁰⁵⁷ Da die Klassifizierung und Verwaltung der Daten Anschaffungs-kosten, Verlagerungskosten und zusätzliche Administrationskosten verursacht, trägt ILM zunächst nicht zu einer Kostensenkung bei. Die Senkung der Speicherplatzkosten genügt nicht, um die zusätzlich entstandenen Anschaffungs- und Betriebskosten auszu-gleichen. Weiterhin ist zu beachten, dass die Daten, trotz der Verwendung der drei Speicherebenen, vorwiegend der Online-Ebene zugeordnet werden. Die jährlich

¹⁰⁵⁶ Die Anzahl der Zugriffe des saisonalen Nutzungsgrads erstreckt sich über sämtliche Intervalle der Klassifizierungsfunktionen. Vgl. hierzu die Funktion des saisonalen Nutzungsgrads Kapitel 5.2.2.1.3.1.4 und die Intervalle der Klassifizierungsfunktionen zur Zuordnung der Daten zu den verschiedenen Speicherebenen Kapitel 5.2.3.1.1.

¹⁰⁵⁷ Vgl. hierzu die Kosten zur Verwaltung der Daten ohne ILM, Kapitel 5.3, Tabelle 5-2., „Speicherplatzkosten“ und die Kosten zur Verwaltung der Daten im ILM, Tabelle 5-9.

schwankenden Zugriffe des saisonalen Nutzungsgrads liegen bei acht von zwölf Monaten, bezüglich der Grundfunktion, bei über dreitausend Zugriffen.¹⁰⁵⁸ Bei einem Anstieg der Zugriffe kommen Speichermedien mit kurzen Zugriffszeiten zum Einsatz. Dies führt zu einer Erhöhung der Verlagerungs- und Administrationskosten. Mit zunehmendem Wachstum der Datenmengen und steigendem Rauschgrad erhöht sich das Kostensenkungspotential. ILM ist für den saisonalen Nutzungsgrad nur wirtschaftlich sinnvoll, wenn große Datenmengen verwaltet werden und der Nutzungsgrad sehr hohe Schwankungen aufweist.¹⁰⁵⁹ Tabelle 5-10 enthält die Kosten, die im zweiten Szenario entstehen.

Rauschgrad [in T]	Anschaffungs- kosten [in T €]	Speicher- platz- kosten [in T €]	Administra- tions- kosten [in T €]	Verlage- rungs- kosten [in T €]	Verlage- rungen [Anzahl]	Gesamt- kosten [in T €]	Kosten- verände- rung [in %]
Datenmenge 1 - 2,856							
0	75	2,0	23	2,6	39	102	59
5	75	2,2	27	3,2	51	107	66
10	75	1,9	22	3,5	62	102	59
20	75	1,4	15	3,6	67	95	47
Datenmenge 10 - 28,561							
0	75	20	227	19	39	340	-47
5	75	22	267	19	51	383	-41
10	75	19	219	18	62	331	-49
20	75	14	148	15	67	252	-61
Datenmenge 100 - 285,61							
0	75	199	2.265	181	39	2.720	-58
5	75	222	2.668	181	51	3.146	-51
10	75	188	2.188	169	62	2.619	-59
20	75	141	1.475	135	67	1.826	-72

Tab. 5-10: Saisonaler Nutzungsgrad, ILM mit Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad

ILM mit Cloud-Ebene trägt hinsichtlich der ersten Datenmenge nicht zu einer Kostensenkung bei. Wirtschaftlichkeit von ILM ist für die weiteren Datenmengen gegeben. Mit zunehmendem Wachstum der Datenmengen und steigendem Rauschgrad erhöht sich das Kostensenkungspotential. Ab einer Anzahl von achttausend Zugriffen werden

¹⁰⁵⁸ Vgl. hierzu die Funktion des saisonalen Nutzungsgrads Kapitel 5.2.2.1.3.1.4.

¹⁰⁵⁹ Vgl. Tabelle 5-9, Datenmenge 100TB bis 285,6TB, Rauschgrad „20000“, Kostenänderung „-5%“. Trotz des hohen Rauschgrads ist ein saisonaler Trend im Nutzungsgrad noch gegeben. Vgl. hierzu Kapitel 5.2.2.1.3.1.4, Abbildung 5-12.

die Daten der Cloud-Ebene zugeordnet.¹⁰⁶⁰ Durch die Erhöhung der Rauschgrade steigt die Anzahl der Verlagerungsvorgänge, die Kosten für die Verlagerungen sinken.¹⁰⁶¹ Die Speicherplatz- und Administrationskosten sinken ebenfalls mit steigendem Rauschgrad.¹⁰⁶²

5.3.5 Sinkender Trend-Saisonaler Nutzungsgrad

Tabelle 5-11 enthält die Kosten für den saisonal variierenden Nutzungsgrad der Informationen mit einer sinkenden Trendkomponente.¹⁰⁶³ ILM ist für den sinkenden trend-saisonalen Nutzungsgrad wirtschaftlich nur sinnvoll, wenn große Datenmengen verwaltet werden und der Nutzungsgrad sehr hohe Schwankungen aufweist.¹⁰⁶⁴ Das Kostensenkungspotential ist gering. Ähnlich wie beim saisonalen Nutzungsgrad ohne Trendkomponente¹⁰⁶⁵ erhöht sich das Kostensenkungspotential mit zunehmendem Wachstum der Daten und steigendem Rauschgrad. Im Vergleich zum saisonalen Nutzungsgrad¹⁰⁶⁶ hat die Verwaltung der Daten mit einem trend-saisonalen Nutzungsgrad aufgrund der sinkenden Trendkomponente¹⁰⁶⁷ dennoch ein höheres Kostensenkungspotential.¹⁰⁶⁸

¹⁰⁶⁰ Vgl. zu den Intervallen der Speicherebenen und Klassifizierungsfunktionen Kapitel 5.2.3.1.1.1.2.

¹⁰⁶¹ Vgl. hierzu Tabelle 5-10, Verlagerungen und Verlagerungskosten. Ab einem Rauschgrad von zehntausend sinken die Verlagerungskosten, da die Daten über den kostengünstigeren Cloud-Import und Cloud-Export verlagert werden. Vgl. hierzu die entsprechenden Kosten für die jeweilige Datenmenge Anhang 5.3.

¹⁰⁶² Da durch eine hohe Anzahl an Zugriffen öfter die Cloud-Ebene verwendet wird. Vgl. hierzu Tabelle 5-10, Speicherplatzkosten, Administrationskosten.

¹⁰⁶³ Vgl. Kapitel 5.2.2.1.3.1.5 zur Beschreibung des Nutzungsgrads.

¹⁰⁶⁴ Vgl. Tabelle 5-11, Datenmenge 100TB bis 285,6TB, Rauschgrad „10000“, Kostenänderung „-2%“ und Rauschgrad „20000“, Kostenänderung „-9%“. Trotz des hohen Rauschgrads ist ein saisonaler Trend im Nutzungsgrad, bezogen auf die 30 Läufe, noch gegeben. Vgl. hierzu Kapitel 5.2.2.1.3.1.5, Abbildung 5-13.

¹⁰⁶⁵ Vgl. Kapitel 5.3.4

¹⁰⁶⁶ Vgl. Kapitel 5.3.4, Tabelle 5-9, „Kostenänderung“

¹⁰⁶⁷ Durch die sinkende Trendkomponente sinkt die Anzahl der Zugriffe, sodass die Daten häufiger auf die kostengünstigeren Ebenen „Nearline“ und „Offline“ verlagert werden. Vgl. zur Klassifizierung der Daten mit Hilfe der Klassifizierungsfunktionen Kapitel 5.2.3.1.1.1.1.

¹⁰⁶⁸ Vgl. hierzu Tabelle 5-11, Kostenänderung sinkender trend-saisonalen Nutzungsgrad und Tabelle 5-9, Kostenänderung saisonaler Nutzungsgrad.

Rauschgrad [in T]	Anschaf- fungs- kosten [in T €]	Speicher- platz- kosten [in T €]	Administra- tions- kosten [in T €]	Verlage- rungs- kosten [in T €]	Verlage- rungen [Anzahl]	Gesamt- kosten [in T €]	Kosten- verände- rung [in %]
Datenmenge 1 - 2,856							
0	60	5,6	77	0,4	4	143	122
5	60	4,8	64	1,7	22	131	103
10	60	4,2	56	2,0	29	122	90
20	60	3,7	53	2,0	32	118	84
Datenmenge 10 - 28,561							
0	60	56	770	4	4	890	38
5	60	48	643	17	22	769	19
10	60	42	561	20	29	683	6
20	60	37	526	20	32	643	0
Datenmenge 100 - 285,61							
0	60	557	7.701	39	4	8.356	30
5	60	484	6.431	173	22	7.149	11
10	60	417	5.614	198	29	6.289	-2
20	60	369	5.255	202	32	5.887	-9

Tab. 5-11: Sinkender Trend-Saisonaler Nutzungsgrad, ILM ohne Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad

Der saisonale Nutzungsgrad unterscheidet sich vom trend-saisonalen Nutzungsgrad darin, dass die jährlichen Schwankungen der Nutzungsgrade beim saisonalen Nutzungsgrad höher sind, als beim trend-saisonalen Nutzungsgrad. Aus diesem Grund werden die Daten des trend-saisonalen Nutzungsgrads ohne Rauschgrad nur vier Mal verlagert, da sich die Anzahl der Zugriffe nicht über sämtliche Intervalle der drei Speicherebenen erstreckt. Mit Erhöhung des Rauschgrads nähern sich die Schwankungen der Zugriffe des trend-saisonalen Nutzungsgrads an die Schwankungen des saisonalen Nutzungsgrads an. Die Anzahl der Verlagerungen sind dann für beide Nutzungsgrade in etwa gleich.¹⁰⁶⁹ In Tabelle 5-12 sind die Kosten enthalten, die im zweiten Szenario, unter Berücksichtigung der Cloud-Ebene, entstehen.

¹⁰⁶⁹ Vgl. hierzu Tabelle 5-11 und 5-9, Anzahl Verlagerungen.

Rauschgrad [in T]	Anschaffungs- kosten [in T €]	Speicher- platz- kosten [in T €]	Administra- tions- kosten [in T €]	Verlage- rungs- kosten [in T €]	Verlage- rungen [Anzahl]	Gesamt- kosten [in T €]	Kosten- verände- rung [in %]
Datenmenge 1 - 2,856							
0	75	3,2	53	0,8	11	132	105
5	75	2,7	36	4,1	63	118	83
10	75	1,9	22	3,8	67	102	59
20	75	1,4	15	3,5	67	95	47
Datenmenge 10 - 28,561							
0	75	32	530	6	11	642	0
5	75	27	360	25	63	487	-24
10	75	19	216	19	67	329	-49
20	75	14	149	15	67	253	-61
Datenmenge 100 - 285,61							
0	75	318	5.298	56	11	5.746	-11
5	75	270	3.602	231	63	4.178	-35
10	75	187	2.157	175	67	2.594	-60
20	75	142	1.485	132	67	1.835	-72

Tab. 5-12: Sinkender Trend-Saisonaler Nutzungsgrad, ILM mit Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad

Die Klassifizierung und Verwaltung der Daten der ersten Datenmenge ist wirtschaftlich nicht sinnvoll. Mit Ausnahme der jährlich wachsenden Datenmenge von 10 Terabyte ohne Rauschgrad,¹⁰⁷⁰ ist die Wirtschaftlichkeit von ILM hinsichtlich der weiteren Datenmengen gegeben. Mit zunehmendem Wachstum der Datenmengen und steigendem Rauschgrad erhöht sich das Kostensenkungspotential. Im Vergleich zum saisonalen Nutzungsgrad¹⁰⁷¹ hat die Verwaltung der Daten mit einem trend-saisonalen Nutzungsgrad aufgrund der sinkenden Trendkomponente¹⁰⁷² im zweiten Szenario ein niedrigeres Kostensenkungspotential.¹⁰⁷³

¹⁰⁷⁰ Das Potential zur Kostensenkung entspricht einem Wert von „-0,0027%“. Vgl. hierzu Tabelle 5-12, Kostenänderung „0%“. Da die Kosten zur Verwaltung der Daten ohne ILM in etwa den Kosten zur Verwaltung der Daten mit ILM entsprechen, ist der Aufwand zur Implementierung eines ILM-Konzepts nicht lohnend. Vgl. hierzu Ott / Wirtschaftlichkeitsanalyse / 523; Kapitel 5.3, Tabelle 5-2, Gesamtkosten, Datenmenge 10TB bis 28,56TB und Tabelle 5-12, Gesamtkosten, Datenmenge 10TB bis 28,56TB, Rauschgrad „0“.

¹⁰⁷¹ Vgl. Kapitel 5.3.4, Tabelle 5-10, „Kostenänderung“

¹⁰⁷² Im zweiten Szenario trägt die Verwendung der Cloud-Ebene maßgeblich zur Senkung der Kosten bei. Durch die sinkende Trendkomponente sinkt die Anzahl der Zugriffe, sodass die Daten häufiger auf die Ebenen „Nearline“ und „Offline“ verlagert werden und weniger auf die Cloud-Ebene.

¹⁰⁷³ Vgl. hierzu Tabelle 5-12, Kostenänderung sinkender trend-saisonalen Nutzungsgrad und Tabelle 5-10, Kostenänderung saisonaler Nutzungsgrad.

5.3.6 Steigender Trend-Saisonaler Nutzungsgrad

Die in Tabelle 5-13 enthaltenen Kosten für den saisonal variierenden Nutzungsgrad mit steigender Trendkomponente¹⁰⁷⁴ sind im ersten Szenario höher, als die Kosten für den saisonal sinkenden Nutzungsgrad.¹⁰⁷⁵ Der Grund hierfür ist, dass die Daten aufgrund der steigenden Trendkomponente ab dem ersten Jahr vorwiegend der kostenintensiven Online-Ebene zugeordnet werden. Die Betriebskosten steigen jährlich um fünf Prozent.¹⁰⁷⁶ Durch die Verwaltung der Daten des steigenden trend-saisonalen Lebenszyklus im ILM entstehen höhere Kosten, als durch die Verwaltung der Daten ohne ILM-Konzept. Die Wirtschaftlichkeit von ILM ist lediglich bei der dritten Datenmenge¹⁰⁷⁷ und einem Rauschgrad von zwanzigtausend gegeben. Durch den starken Rauschgrad wird die Grundfunktion des steigenden trend-saisonalen Nutzungsgrads¹⁰⁷⁸ derart verändert, dass die Zugriffe bzw. Nutzungsgrade zusätzlich den Intervallen der Speicherebenen „Near-line“ und „Offline“ zugeordnet werden können. Die Kostensenkung beträgt, im Vergleich zur Verwaltung der Daten ohne ILM, drei Prozent. ILM ist ohne Cloud-Ebene nur wirtschaftlich, wenn große Datenmengen verwaltet werden und wenn die Zugriffe sehr hohen Schwankungen unterliegen.

¹⁰⁷⁴ Vgl. Kapitel 5.2.2.1.3.1.6 zur Beschreibung des Nutzungsgrads.

¹⁰⁷⁵ Vgl. Kapitel 5.3.5

¹⁰⁷⁶ Vgl. hierzu Anhang 5.3, Tabelle A5-3.

¹⁰⁷⁷ Vgl. Tabelle 5-13, Datenmenge 100TB bis 285,6TB, Rauschgrad „20000“

¹⁰⁷⁸ Vgl. Kapitel 5.2.2.1.3.1.6

Rauschgrad [in T]	Anschaf- fungs- kosten [in T €]	Speicher- platz- kosten [in T €]	Administra- tions- kosten [in T €]	Verlage- rungs- kosten [in T €]	Verlage- rungen [Anzahl]	Gesamt- kosten [in T €]	Kosten- verände- rung [in %]
Datenmenge 1 - 2,856							
0	60	5,4	84	0,2	4	150	132
5	60	4,7	75	1,2	22	141	119
10	60	4,0	64	1,7	29	130	102
20	60	3,7	56	2,0	32	122	89
Datenmenge 10 - 28,561							
0	60	54	840	2	4	956	48
5	60	47	749	12	22	868	35
10	60	40	643	17	29	761	18
20	60	37	563	20	32	679	5
Datenmenge 100 - 285,61							
0	60	542	8.401	18	4	9.021	40
5	60	468	7.491	122	22	8.141	26
10	60	404	6.434	173	29	7.071	10
20	60	366	5.627	199	32	6.252	-3

Tab. 5-13: Steigender Trend-Saisonaler Nutzungsgrad, ILM ohne Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad

In Tabelle 5-14 sind die Kosten für den steigenden trend-saisonalen Nutzungsgrad enthalten, die bei der automatisierten Verwaltung der Daten im zweiten Szenario entstehen. Der Einsatz von ILM mit Cloud-Ebene ist bezüglich der ersten Datenmenge nicht wirtschaftlich.¹⁰⁷⁹ Wirtschaftlichkeit von ILM ist für die weiteren Datenmengen gegeben.¹⁰⁸⁰ Mit zunehmendem Wachstum der Daten und steigendem Rauschgrad erhöht sich das Kostensenkungspotential.¹⁰⁸¹

¹⁰⁷⁹ Vgl. hierzu Tabelle 5-14, Kostenänderung „73%“, „74%“, „60%“, „46%“.

¹⁰⁸⁰ Vgl. hierzu Kostenänderung für die jeweiligen Datenmengen und Rauschgrade Tabelle 5-14.

¹⁰⁸¹ Vgl. Tabelle 5-14, Kostenänderung der drei Datenmengen

Rauschgrad [in T]	Anschaffungs- kosten [in T €]	Speicher- platz- kosten [in T €]	Administra- tions- kosten [in T €]	Verlage- rungs- kosten [in T €]	Verlage- rungen [Anzahl]	Gesamt- kosten [in T €]	Kosten- verände- rung [in %]
Datenmenge 1 - 2,856							
0	75	3,5	33	0,5	10	112	73
5	75	2,7	31	3,8	60	112	74
10	75	2,0	22	3,8	67	103	60
20	75	1,4	14	3,4	66	94	46
Datenmenge 10 - 28,561							
0	75	35	327	4	10	441	-32
5	75	27	308	21	60	430	-33
10	75	20	222	19	67	336	-48
20	75	14	144	14	66	247	-62
Datenmenge 100 - 285,61							
0	75	352	3.269	34	10	3.729	-42
5	75	267	3.075	198	60	3.615	-44
10	75	196	2.224	178	67	2.672	-59
20	75	139	1.441	126	66	1.781	-72

Tab. 5-14: Steigender Trend-Saisonaler Nutzungsgrad, ILM mit Cloud-Ebene, 5 Jahre, 30 Läufe je Rauschgrad

5.3.7 Zusammenfassung der Wirtschaftlichkeitsanalyse

Unter Berücksichtigung der getroffenen Annahmen zu den Klassifizierungsfunktionen,¹⁰⁸² Verlagerungsfunktionen,¹⁰⁸³ Kostenfunktionen¹⁰⁸⁴ und Betriebskosten¹⁰⁸⁵ werden in diesem Kapitel die wesentlichen Ergebnisse der Wirtschaftlichkeitsanalyse zusammengefasst. Die Verwaltung von kleinen Datenmengen im ILM ist wirtschaftlich nicht sinnvoll.¹⁰⁸⁶ Mit Ausnahme des sinkenden Nutzungsgrads gilt für alle untersuchten Nutzungsgrade, dass sich das Wachstum der Daten¹⁰⁸⁷ und die Variation der Zugriffe im Zeitverlauf¹⁰⁸⁸ positiv auf das Kostensenkungspotential auswirken.¹⁰⁸⁹ Im ersten

¹⁰⁸² Vgl. Kapitel 5.2.3.1

¹⁰⁸³ Vgl. Kapitel 5.2.3.2

¹⁰⁸⁴ Vgl. Kapitel 5.2.3.3.2

¹⁰⁸⁵ Vgl. Kapitel 5.2.3.3.1, Anhang 5.3

¹⁰⁸⁶ Vgl. hierzu Kapitel 5.3.1 bis 5.3.6. Dies gilt für die beiden betrachteten ILM-Szenarien. Es werden Speicherplatzkosten gesenkt. Wirtschaftlichkeit ist aufgrund der Anschaffungs- und Betriebskosten im ILM dennoch nicht gegeben, die Daten sollten nicht mit Hilfe von ILM verwaltet werden.

¹⁰⁸⁷ Das Wachstum der Daten beträgt jährlich dreißig Prozent. Vgl. hierzu Kapitel 5.2.2.1.

¹⁰⁸⁸ Simuliert durch die verschiedenen Rauschgrade. Vgl. hierzu Kapitel 5.2.2.1.3.

¹⁰⁸⁹ Vgl. hierzu Kapitel 5.3.

Szenario ist die Wirtschaftlichkeit von ILM beim sinkenden Nutzungsgrad für die jährlich wachsenden Datenmengen von 10 Terabyte und 100 Terabyte für sämtliche Rauschgrade gegeben.¹⁰⁹⁰ ILM ohne Cloud-Ebene ist hinsichtlich des steigenden, saisonalen, sinkend-trend saisonalen und steigend-trend saisonalen Nutzungsgrads wirtschaftlich nur sinnvoll, wenn große Datenmengen¹⁰⁹¹ verwaltet werden und wenn der jeweilige Nutzungsgrad sehr stark variiert.¹⁰⁹² Die Verwaltung der Daten des betrachteten konstanten Nutzungsgrads führt nicht zu Kostensenkungen, da die Daten durchgängig der Online-Ebene zugehörig sind. Ob die Verwaltung der Daten eines konstanten Nutzungsgrads wirtschaftlich ist, hängt von der Anzahl der Zugriffe ab. Finden bspw. konstant wenig oder keine Zugriffe statt, so ist die Verwaltung der Daten im ILM wirtschaftlich, da diese über die Dauer des Lebenszyklus kostengünstigeren Speicherebenen zugeordnet werden können. Denkbar ist ebenfalls, dass die Zugriffe konstant sehr hoch sind und die Daten ausschließlich auf der Cloud-Ebene verwaltet werden können. Dies würde ebenfalls zu einer wirtschaftlichen Verwaltung des konstanten Nutzungsgrads beitragen. Es wird empfohlen, insbesondere Daten mit einem sinkenden Nutzungsgrad mit Hilfe von ILM zu verwalten.¹⁰⁹³ Die zusätzliche Berücksichtigung der Cloud-Ebene führt im zweiten Szenario bei allen Nutzungsgraden zu einer Erhöhung der Wirtschaftlichkeit.¹⁰⁹⁴ Die Nutzung der Cloud-Ebene eignet sich vor allem für große Datenmengen mit hohen Zugriffsschwankungen.¹⁰⁹⁵ Die Verlagerungskosten sind in beiden Szenarien nicht ausschlaggebend dafür, inwieweit die automatisierte Verwaltung der Daten zu einer Kostensenkung führt.¹⁰⁹⁶

¹⁰⁹⁰ Vgl. hierzu Kapitel 5.3.1. Aufgrund der Abnahme der Anzahl der Zugriffe ab dem ersten Jahr (Vgl. Kapitel 5.2.2.1.3.1.1) werden die Daten trotz der Rauschgrade vorwiegend den kostengünstigeren Speicherebenen „Nearline“ und „Offline“ zugeordnet.

¹⁰⁹¹ Vgl. hierzu die Datenmengen von 10TB bis 28,56TB bzw. 100TB bis 285,6TB in Anhang 1.

¹⁰⁹² Vgl. hierzu Kapitel 5.3.2, Kapitel 5.3.4, Kapitel 5.3.5 und Kapitel 5.3.6.

¹⁰⁹³ Vgl. Kapitel 5.3.1

¹⁰⁹⁴ Bezüglich der jährlich wachsenden Datenmengen von 10 Terabyte und 100 Terabyte. Vgl. hierzu Kapitel 5.3.1 bis Kapitel 5.3.6. Ausgenommen hiervon ist der konstante Nutzungsgrad mit einem Rauschgrad von null, da die Daten hier ausschließlich der Online-Ebene zugeordnet werden. Vgl. hierzu Kapitel 5.3.3.

¹⁰⁹⁵ Vgl. hierzu Kapitel 5.3.1 bis Kapitel 5.3.6. Bei einem schnellen Anstieg der Zugriffe kann die benötigte Speicherplatzkapazität durch die Nutzung der IT-Infrastruktur des IaaS-Anbieters flexibel erweitert werden. Vgl. hierzu Kapitel 3.2.2.

¹⁰⁹⁶ Vgl. hierzu die Verlagerungskosten Kapitel 5.3.1 bis Kapitel 5.3.6. Im Hinblick auf die Wirtschaftlichkeit sind insbesondere die Speicherplatzkosten und die Administrationskosten für die Verwaltung der Daten entscheidend.

5.4 Evaluation der Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM

Ziel der Evaluation ist es, die zur Lösung eines bestimmten Problems definierten Objekte auf der Grundlage eines Systems von Kriterien zielbezogen zu beurteilen.¹⁰⁹⁷ Bei der Evaluation kann zwischen der künstlichen und naturalistischen bzw. realitätsnahen Evaluation unterschieden werden.¹⁰⁹⁸ Zu den künstlichen Evaluationsmethoden gehören bspw. Laborexperimente, Simulationen, mathematische Methoden oder Merkmalsvergleiche.¹⁰⁹⁹ Im Gegensatz dazu erfolgt die Evaluation von Artefakten bei der naturalistischen Evaluation in einer realen Umgebung, bspw. in Form von Umfragen oder Fallstudien.¹¹⁰⁰ Durch künstliche Evaluation lassen sich Aussagen und Erkenntnisse erarbeiten, die auf die Wirklichkeit übertragbar sind,¹¹⁰¹ allerdings handelt es sich bei künstlichen Evaluationsmethoden um Näherungsverfahren, bei denen die spezifischen Anforderungen einer einzelnen Unternehmung oder das Wissen von Anwendern und Administratoren nicht berücksichtigt werden.¹¹⁰² Der naturalistischen Evaluation wird an Kritik entgegengebracht, dass die Ergebnisse der Untersuchungen von Fehlinterpretationen beeinflusst werden können und sich lediglich auf spezifische Gegebenheiten einzelner Organisationen beziehen.¹¹⁰³ In der vorliegenden Arbeit wird die Wirtschaftlichkeitsanalyse im ILM mit Hilfe der Methoden der Simulation¹¹⁰⁴ und des Merkmalsvergleichs¹¹⁰⁵ demonstriert und evaluiert.¹¹⁰⁶ Ziel der Arbeit ist die Wirtschaftlichkeitsanalyse der Verwaltung von unstrukturierten Daten. Hierzu werden Lebenszyklen von Daten untersucht, die typischerweise in der Realität vorkommen. Die Lebenszyklen werden nicht

¹⁰⁹⁷ Vgl. Heinrich / Evaluationsforschung / 9; Mertens / Evaluation / 47

¹⁰⁹⁸ Vgl. Venable, Pries-Heje, Baskerville / Framework for Evaluation / 428-429; Hevner et al. / Design Science in IS / 75-105

¹⁰⁹⁹ Vgl. Nunamaker, Chen, Purdin / Systems Development / 95

¹¹⁰⁰ Vgl. Sun, Kantor / Information System Evaluation / 614-628

¹¹⁰¹ Vgl. VDI / Simulation / 2

¹¹⁰² Vgl. Sun, Kantor / Information System Evaluation / 614-628; Scheuermann / Simulationen / 3; Bossel / Simulation komplexer Systeme / 109

¹¹⁰³ Vgl. Frank / Research Methods in IS / 5-6; Venable, Pries-Heje, Baskerville / Framework for Evaluation / 429; Satyanarayanan / Study Lifetimes / 99; Venable / Design Science Research / 117

¹¹⁰⁴ Vgl. Kapitel 5.2; Kapitel 5.3

¹¹⁰⁵ Vgl. Kapitel 5.2.1

¹¹⁰⁶ Die Demonstration ist die Vorstufe der Evaluation und kann als ein Teil der Evaluation betrachtet werden. Vgl. hierzu Venable, Pries-Heje, Baskerville / Framework for Evaluation / 428.

aus einer einzelnen Organisation oder Unternehmung erhoben. Vielmehr findet der Rückgriff auf Forschungsergebnisse zu Verhaltensstudien zur Nutzung von Informationen statt.¹¹⁰⁷ Es werden Aussagen zur wirtschaftlich sinnvollen Nutzung von unstrukturierten Daten erarbeitet, die unabhängig von bestimmten organisatorischen oder technischen Gegebenheiten einer einzelnen Unternehmung sind. Daraus können bspw. Handlungsempfehlungen zur kosteneffizienten Verwaltung von unstrukturierten Daten abgeleitet werden und Aussagen dazu getroffen werden, inwieweit eine Einführung und Umsetzung von ILM in der Praxis aus Wirtschaftlichkeitsgesichtspunkten sinnvoll ist.¹¹⁰⁸

Die Demonstration ist dazu geeignet, den Wertbeitrag des konstruierten und implementierten Verfahrens zur Wirtschaftlichkeitsanalyse¹¹⁰⁹ im Hinblick auf die Erfüllung der Zielstellung der Arbeit¹¹¹⁰ zu untersuchen.¹¹¹¹ Die Demonstration der Wirtschaftlichkeitsanalyse sowie die Auswertung der entsprechenden Ergebnisse erfolgte in Kapitel 5.3.¹¹¹² Es wurde eine Wirtschaftlichkeitsanalyse der automatisierten Verwaltung unstrukturierter Daten im ILM für verschiedene Lebenszyklen durchgeführt.¹¹¹³ Neben der Verwaltung unstrukturierter Daten auf unternehmenseigenen Speichermedien wurde zusätzlich das Konzept des Cloud Computings¹¹¹⁴ zur Verwaltung der Daten mit Hilfe von Diensten über das Internet¹¹¹⁵ betrachtet.¹¹¹⁶ Das definierte Ziel der Arbeit wurde erreicht.

In Kapitel 5.2.1 wurden Anforderungen an die Wirtschaftlichkeitsanalyse im ILM definiert. Im Folgenden wird mit Hilfe eines Merkmalsvergleichs überprüft, inwieweit die

¹¹⁰⁷ Vgl. Kapitel 1.3

¹¹⁰⁸ Vgl. hierzu Kapitel 5.3; Frank / Formale Sprachen / 142-143; Satyanarayanan / Study Lifetimes / 99; Turczyk / ILM / 87; Venable / Design Science Research / 117; Venable / Framework for Design Science Research Activities / 186

¹¹⁰⁹ Vgl. Kapitel 5.2

¹¹¹⁰ Vgl. Kapitel 1.2

¹¹¹¹ Vgl. Peffers et al. / Research Methodology / 55; Venable, Pries-Heje, Baskerville / Framework for Evaluation / 428; Walls, Widmeyer, El Sawy / Design Theory / 36-59; Vgl. Kapitel 5.3

¹¹¹² Hierbei handelt es sich um eine nachträgliche bzw. „ex post Evaluation“, bei der die Ergebnisse der Simulation ausgewertet werden. Vgl. hierzu Venable, Pries-Heje, Baskerville / Framework for Evaluation / 432-433. Vgl. zur Unterscheidung von „ex post Evaluation“ und „ex ante Evaluation“ Venable, Pries-Heje, Baskerville / Framework for Evaluation / 429-430.

¹¹¹³ Vgl. zur Zusammenfassung der Ergebnisse Kapitel 5.3.7

¹¹¹⁴ Vgl. Kapitel 3.1

¹¹¹⁵ Vgl. Kapitel 3.2.1

¹¹¹⁶ Vgl. Kapitel 5.3

Anforderungen hinsichtlich der durchgeführten bzw. demonstrierten Wirtschaftlichkeitsanalyse erfüllt wurden:

- Analyse der Kosten- und Nutzenstruktur: Die Kosten- und Nutzenstruktur des ILM wurde analysiert.¹¹¹⁷ Relevante Kostengrößen sind die Anschaffungs- und Betriebskosten, die für die Implementierung eines ILM-Konzepts zur automatisierten Klassifizierung von Informationen und für die Verwaltung der Informationen in den unterschiedlichen Speicherebenen entstehen.¹¹¹⁸ Der Nutzen resultiert aus den erzielten Kostensenkungen.¹¹¹⁹
- Analyse der Beziehungszusammenhänge von Kosten und Nutzen: Es wurde untersucht, welchen Einfluss die Verwaltung unterschiedlich großer Datenmengen mit variierenden Zugriffshäufigkeiten und Nutzungsgraden auf die Wirtschaftlichkeit des ILM hat. Mit Hilfe der Kostenvergleichsrechnung¹¹²⁰ wurden die Kosten, die bei der Verwaltung der Daten ohne und mit ILM entstehen, gegenübergestellt und miteinander verglichen.¹¹²¹
- Automatisierung der Wirtschaftlichkeitsanalyse: Die Analyse der Wirtschaftlichkeit¹¹²² der Klassifizierung der Daten der jeweiligen Nutzungsgrade¹¹²³ wurde automatisiert durchgeführt.¹¹²⁴ Hierzu wurden Klassifizierungs-, Verlagerungs- und Kostenfunktionen implementiert.¹¹²⁵ Es wurden zwei Szenarien mit

¹¹¹⁷ Vgl. Kapitel 5.2.3.3.1

¹¹¹⁸ Relevante Kosten sind bspw. Beratungs- und Implementierungskosten (Vgl. Anhang 5.1), Speicher- und Administrationskosten, Kosten für die Erhebung der Nutzungsgrade unstrukturierter Daten, Schulungskosten, Verlagerungs- bzw. Migrationskosten, Kosten für die Datenübertragung und Kosten, die bei der Nutzung von Public-Cloud-Diensten anfallen. Vgl. hierzu Anhang 5.3.

¹¹¹⁹ Vgl. Kapitel 5.3

¹¹²⁰ Vgl. Kapitel 4.1

¹¹²¹ Vgl. Kapitel 5.3

¹¹²² Vgl. hierzu die Anforderung der Kostenreduktion und Wirtschaftlichkeit der automatisierten Klassifizierung Kapitel 5.1.3.2.5.

¹¹²³ Vgl. hierzu die Anforderung zur Verwendung des Nutzungsgrads als Klassifizierungskriterium Kapitel 5.1.3.2.1

¹¹²⁴ Vgl. hierzu die Anforderungen an Konzepte zur automatisierten Klassifizierung unstrukturierter Daten Kapitel 5.1.3.2.

¹¹²⁵ Planungsaufwände zur Durchführung der Wirtschaftlichkeitsanalyse können somit reduziert werden. Vgl. hierzu Becker / Wirtschaftswissen / 180. Vgl. hierzu die Anforderung zur Implementierung geeigneter Klassifizierungsfunktionen Kapitel 5.1.3.2.2.

verschiedenen Nutzungsgraden, Klassifizierungsfunktionen, Speicherebenen und Datenmengen¹¹²⁶ simuliert, um die jeweiligen Kosten und Kostensenkungspotentiale zu ermitteln.

- Berücksichtigung des Cloud Computings zur Verwaltung der Informationen: Zusätzlich zu den unternehmensinternen Speichermedien wurde das Cloud Computing zur wirtschaftlichen Verwaltung der Informationen betrachtet.¹¹²⁷
- Untersuchung der Ergebnisse: Die Ergebnisse der Wirtschaftlichkeitsanalyse wurden untersucht, um die Wirtschaftlichkeit der automatisierten Verwaltung unstrukturierter Daten im ILM beurteilen zu können.¹¹²⁸
- Berücksichtigung des Anwender- und Administratorenwissens: Eine Befragung von Anwendern, Administratoren oder Experten bezüglich der Nutzung und Klassifizierung der Informationen und den relevanten rechtlichen Rahmenbedingungen¹¹²⁹ wurde nicht durchgeführt.¹¹³⁰

Die Wirtschaftlichkeit von ILM hängt von der jeweiligen Nutzung der Informationen, den Kosten für die verwendeten Speichermedien und Speicherebenen, den Datenmengen sowie den definierten Klassifizierungsfunktionen¹¹³¹ ab. Die Ergebnisse¹¹³² der Wirtschaftlichkeitsanalyse¹¹³³ variieren entsprechend der jeweiligen Gegebenheiten. Eine wichtige Voraussetzung für die Wirtschaftlichkeitsanalyse unstrukturierter Daten im ILM ist die Extraktion von Metadaten. Die Nutzungsgrade von unstrukturierten Daten, wie bspw. PDF- oder Microsoft-Office-Dateien, werden in Form von XML-

¹¹²⁶ Vgl. hierzu die Anforderung zur Berücksichtigung großer Datenmengen Kapitel 5.1.3.2.6.

¹¹²⁷ Vgl. Kapitel 5.2.3.1.1.1.2; Kapitel 5.2.3.3.2.1.2

¹¹²⁸ Vgl. Kapitel 4.1; Kapitel 5.3

¹¹²⁹ Vgl. hierzu die Anforderung zur Berücksichtigung des Anwender- und Administratorenwissens und rechtlicher Rahmenbedingungen Kapitel 5.1.3.2.3 und 5.1.3.2.4.

¹¹³⁰ Die Analysen wurden nicht bezogen auf eine einzelne Unternehmung durchgeführt. Vgl. hierzu Kapitel 1.3.

¹¹³¹ Die Klassifizierungsfunktionen sollten so gestaltet werden, dass die Zugriffsintervalle für kostengünstige Speicherebenen, im Gegensatz zu den Intervallen der kostenintensiven Speicherebenen, einen möglichst großen Wertebereich annehmen. Allerdings sollte beachtet werden, dass die kostengünstigeren unternehmensinternen Speichermedien längere Zugriffszeiten aufweisen.

¹¹³² Vgl. Kapitel 5.3

¹¹³³ Die Wirtschaftlichkeitsanalyse basiert auf der prototypischen Implementierung. Vgl. hierzu Kapitel 5.2.3.

basierten Daten¹¹³⁴ innerhalb der Dateien systemseitig gespeichert.¹¹³⁵ Zum Auslesen der Daten können entsprechende Softwarewerkzeuge verwendet werden.¹¹³⁶ Sollte das Auslesen bzw. die Verwendung dieser Daten nicht möglich sein,¹¹³⁷ kann die Klassifizierung der Informationen nicht automatisiert anhand der Nutzungsgrade erfolgen.¹¹³⁸ Dieser Sachverhalt würde sich auf die Wirtschaftlichkeit der Verwaltung der Daten auswirken, da die Kosten für die Klassifizierung steigen würden. In dieser Arbeit wurden den Untersuchungen zur Wirtschaftlichkeitsanalyse im ILM marktübliche Speicher- und Administrationskosten sowie typische in der Realität vorkommende Lebenszyklen zugrunde gelegt.¹¹³⁹ Es wurde gezeigt, dass ILM zur wirtschaftlich sinnvollen Verwaltung unstrukturierter Daten beitragen kann.¹¹⁴⁰ Weiterhin wurde erörtert, unter welchen Rahmenbedingungen die Verwaltung unstrukturierter Daten wirtschaftlich ist.

6 Schlussbetrachtungen

6.1 Zusammenfassung und kritische Würdigung der Arbeit

Informationen müssen einerseits für betriebliche Aufgaben schnell verfügbar sein. Andererseits zwingt ein steigender Kostendruck viele Unternehmen dazu, Informationen auf möglichst kostengünstigen Speichern mit langen Zugriffszeiten vorzuhalten. ILM hilft zu entscheiden, welche Informationen auf teuren und schnellen Speichermedien verbleiben müssen und welche auf langsame, aber kostengünstige Medien verlagert werden können.¹¹⁴¹ ILM unterstützt die Analyse und Klassifizierung der betrieblichen Informationen,¹¹⁴² die Ermittlung der jeweils angemessenen Speichermedien und Speicherebenen sowie die eventuelle Verlagerung der Informationen.¹¹⁴³ Die Untersuchun-

¹¹³⁴ Vgl. Kapitel 2.4.2

¹¹³⁵ Vgl. Wootton / Developing Quality Metadata / 200-201; Witten et al. / Documents and Metadata / 389-405; Baker / Adobe Acrobat / 194-195; Antonovich / Office User's Guide / 222-231

¹¹³⁶ Wie bspw. das „Metadata Extraction Tool“ (Vgl. NLNZ / Metadata Extraction Tool / .

¹¹³⁷ Bspw. weil ein bestimmtes Dokumentformat die systemseitige Speicherung von Nutzungsdaten nicht unterstützt. Oder weil bspw. aufgrund unternehmensspezifischer rechtlicher oder sicherheitstechnischer Anforderungen das Auslesen dieser Daten nicht erlaubt ist.

¹¹³⁸ Dies könnte dazu führen, dass eine automatisierte Unterstützung der Klassifizierung nicht möglich ist und die Klassifizierung der Informationen bspw. manuell durchgeführt werden muss.

¹¹³⁹ Vgl. Kapitel 1.3, Anhang 5

¹¹⁴⁰ Vgl. Kapitel 5.3

¹¹⁴¹ Vgl. Kapitel 2.4.2; Kapitel 5.1.2

¹¹⁴² Vgl. Kapitel 5.2.2.1.3.1

¹¹⁴³ Vgl. Kapitel 5.2.3.1; Kapitel 5.2.3.2

gen in dieser Arbeit haben ergeben, dass insbesondere die automatisierte Unterstützung der Informationsklassifizierung erst wenig entwickelt ist.¹¹⁴⁴ Die vorliegende Arbeit ist ein Schritt zur Verbesserung und Weiterentwicklung der automatisierten Verwaltung unstrukturierter Daten im ILM. Die rechtlichen Anforderungen sind sehr komplex und werden vom Gesetzgeber häufig verändert bzw. angepasst.¹¹⁴⁵ Weder die in Kapitel 5.1.3.1 vorgestellten Konzepte zur automatisierten Klassifizierung noch die in dieser Arbeit implementierten Klassifizierungsfunktionen¹¹⁴⁶ berücksichtigen rechtliche Anforderungen. Dieses Gebiet ist ein offenes Forschungsthema.¹¹⁴⁷

Die Kosten für die Durchführung des ILM dürfen die durch das ILM erreichbaren Kostensenkungen nicht übersteigen.¹¹⁴⁸ Zur Analyse des Kosten- und Nutzenverhältnisses, das bei der Implementierung und Anwendung von ILM entsteht, wurde in dieser Arbeit eine Wirtschaftlichkeitsanalyse durchgeführt.¹¹⁴⁹ Hierzu wurden Klassifizierungs-, Verlagerungs- und Kostenfunktionen implementiert, um die Entstehung von Kosten und Kostensenkungspotentialen¹¹⁵⁰ für verschiedene Nutzungsgrade zu simulieren und auszuwerten. Es wurde untersucht, welchen Einfluss die Verwaltung unterschiedlich großer Datenmengen mit variierenden Zugriffshäufigkeiten und Nutzungsgraden auf die Wirtschaftlichkeit des ILM hat.¹¹⁵¹ Die Untersuchungen haben gezeigt, dass die Verwaltung von kleinen Datenmengen im ILM wirtschaftlich nicht sinnvoll ist.¹¹⁵² Mit Ausnahme des sinkenden Nutzungsgrads¹¹⁵³ gilt für alle untersuchten Nutzungsgrade, dass sich das Wachstum der Daten¹¹⁵⁴ und die Variation der Zugriffe im Zeitverlauf¹¹⁵⁵ positiv auf

¹¹⁴⁴ Vgl. Kapitel 5.1.3.1

¹¹⁴⁵ Vgl. Kaarst-Brown, Kelly / IT-Governance /

¹¹⁴⁶ Vgl. Kapitel 5.2.3.1

¹¹⁴⁷ Vgl. Abd-El-Malek et al. / Cluster-based storage / 59-72; Mont, Beato / Obligation Policies / 51-55; Hadjiantonis, Pavlou / Self-Management / 257-258; Krishnan / Big Data / 113-114; Shen et al. / Data Protection in the Cloud / 137-139

¹¹⁴⁸ Vgl. Kapitel 5.3

¹¹⁴⁹ Vgl. Kapitel 5

¹¹⁵⁰ In dieser Arbeit wurden ausschließlich monetäre Nutzengrößen bzw. messbarer Nutzen untersucht. Vgl. hierzu Kapitel 4.3.

¹¹⁵¹ Vgl. Kapitel 5.3.7

¹¹⁵² Vgl. Kapitel 5.3.1 bis 5.3.6

¹¹⁵³ Im ersten Szenario ist die Wirtschaftlichkeit von ILM beim sinkenden Nutzungsgrad für die jährlich wachsenden Datenmengen von 10 Terabyte und 100 Terabyte für sämtliche Rauschgrade gegeben. Vgl. hierzu Kapitel 5.3.1.

¹¹⁵⁴ Vgl. Kapitel 5.2.2.1

das Kostensenkungspotential auswirken.¹¹⁵⁶ ILM ist im ersten Szenario¹¹⁵⁷ hinsichtlich des steigenden, saisonalen, sinkend-trend saisonalen und steigend-trend saisonalen Nutzungsgrads wirtschaftlich nur sinnvoll, wenn große Datenmengen¹¹⁵⁸ verwaltet werden und wenn der jeweilige Nutzungsgrad sehr stark variiert.¹¹⁵⁹ Im zweiten Szenario wurde untersucht, welche Auswirkungen das Cloud Computing auf die Wirtschaftlichkeit der Verwaltung unstrukturierter Daten im ILM hat.¹¹⁶⁰ Die Berücksichtigung der Cloud-Ebene führt im zweiten Szenario bei allen Nutzungsgraden zu einer Erhöhung der Wirtschaftlichkeit.¹¹⁶¹ Die Nutzung der Cloud-Ebene eignet sich vor allem für große Datenmengen mit hohen Zugriffsschwankungen.¹¹⁶²

Kritisch zu betrachten ist, dass die automatisierte Klassifizierung der Informationen in dieser Arbeit ausschließlich anhand von Nutzungsgraden und Datenmengen erfolgt. Die implementierten Klassifizierungsfunktionen schlagen eine geeignete Speicherebene vor, auf die die Daten verlagert werden sollen. Wenn die Klassifizierung von Informationen allein auf Basis von Metadaten und technischen Kriterien erfolgt, so müssen die Ergebnisse der Klassifizierung in der Praxis von Anwendern und Administratoren überprüft werden, bevor eine Verlagerung stattfindet. Die Verlagerung der Informationen ist in Abhängigkeit von der verwendeten Speicherebene mit Zeit- und Kostenaufwänden verbunden.¹¹⁶³ Bei stark schwankenden Nutzungsgraden können bspw. häufige Verlagerungsvorgänge der Informationen zwischen den Speicherebenen notwendig werden. Weiterhin ist zu beachten, dass für die Durchführung der automatisierten Klassifizierung die entsprechenden Metadaten und Nutzungsgrade für unstrukturierte Daten vorhanden sein müssen. In dieser Arbeit wurden Daten für Lebenszyklen generiert, die typischerweise in der Praxis vorkommen.¹¹⁶⁴ Für die Umsetzung der automatisierten

¹¹⁵⁵ Vgl. Kapitel 5.2.2.1.3

¹¹⁵⁶ Vgl. hierzu Kapitel 5.3.7. Dennoch führt die Verwaltung der Daten des konstanten Nutzungsgrads nicht zu Kostensenkungen. Vgl. hierzu Kapitel 5.3.3.

¹¹⁵⁷ Vgl. Kapitel 5.2.3.1.1.1.1

¹¹⁵⁸ Vgl. hierzu die Datenmengen von 10TB bis 28,56TB bzw. 100TB bis 285,6TB in Anhang 1.

¹¹⁵⁹ Vgl. hierzu Kapitel 5.3.2, Kapitel 5.3.4, Kapitel 5.3.5 und Kapitel 5.3.6.

¹¹⁶⁰ Vgl. Kapitel 5.2.3.1.1.1.2

¹¹⁶¹ Vgl. hierzu Kapitel 5.3.1 bis Kapitel 5.3.6.

¹¹⁶² Vgl. hierzu Kapitel 5.3.1 bis Kapitel 5.3.6.

¹¹⁶³ Vgl. Kapitel 3.4.1

¹¹⁶⁴ Vgl. Kapitel 1.3

Klassifizierung in einem Unternehmen müssen die Metadaten von unstrukturierten Daten entsprechend erhoben werden.¹¹⁶⁵ Dies setzt voraus, dass Metadaten vorhanden sind und von den verwendeten Betriebssystemen und Dokumentenverwaltungssystemen für die jeweiligen Dateiformate protokolliert und abgespeichert werden. Ist dies nicht der Fall, müssen Mechanismen implementiert werden, die die Nutzung der Informationen im Zeitverlauf protokollieren. Wenn keine Metadaten vorhanden sind und auch nicht protokolliert werden können, so kann die automatisierte Klassifizierung anhand des Nutzungsgrads nicht erfolgen. In diesem Fall wäre zwar immer noch eine Klassifizierung der Informationen bspw. durch Anwender und Administratoren möglich. Allerdings würde dies die Wirtschaftlichkeit des ILM negativ beeinflussen, da der Aufwand für die Klassifizierung der Informationen deutlich höher wäre.

Durch Simulationsläufe bzw. Experimente lassen sich zwar Aussagen und Erkenntnisse erarbeiten, die auf die Wirklichkeit übertragbar sind,¹¹⁶⁶ allerdings stellt die Verwendung der Methode der Simulationstechnik zugleich eine Grenze der Wirtschaftlichkeitsanalyse¹¹⁶⁷ dar.¹¹⁶⁸ Bei der in dieser Arbeit durchgeführten Simulation handelt es sich um ein Näherungsverfahren, bei dem die spezifischen Anforderungen einer einzelnen Unternehmung und das Wissen von Anwendern und Administratoren nicht berücksichtigt werden.¹¹⁶⁹ Die gewonnenen Erkenntnisse aus der durchgeführten Wirtschaftlichkeitsanalyse¹¹⁷⁰ liefern dennoch wichtige Hinweise zur wirtschaftlich sinnvollen Klassifizierung unstrukturierter Daten im ILM.¹¹⁷¹ Die Wirtschaftlichkeit des ILM lässt sich durch eine Anpassung bzw. Erweiterung der Klassifizierungsfunktionen, über die in Kapitel 5.3 erörterten Kostensenkungspotentiale hinaus, weiter erhöhen. Denkbar wäre, die Klassifizierungsfunktionen so zu gestalten, dass die Zugriffsintervalle für kosten-

¹¹⁶⁵ Vgl. Kapitel 5.2.2.1.2 zur Verwendung von Softwarewerkzeugen zur Analyse der Metadaten von Dokumenten.

¹¹⁶⁶ Vgl. VDI / Simulation / 2

¹¹⁶⁷ Vgl. Kapitel 5.2; Kapitel 5.3

¹¹⁶⁸ Die Simulationen wurden mit Hilfe von Microsoft Excel durchgeführt, da dies für die Experimente in dieser Arbeit ausreichend war. Für komplexere Simulationen sollte ggf. auf ein geeignetes Simulationswerkzeug zurückgegriffen werden.

¹¹⁶⁹ Vgl. zu den Grenzen der Simulationstechnik Scheuermann / Simulationen / 3; Bossel / Simulation komplexer Systeme / 109.

¹¹⁷⁰ Vgl. Kapitel 5.3

¹¹⁷¹ Für bspw. die Implementierung von Klassifizierungsfunktionen in der Praxis sind diese entsprechend den jeweiligen Unternehmensanforderungen anzupassen.

günstige Speicherebenen, im Gegensatz zu den Intervallen der kostenintensiven Speicherebenen, einen möglichst großen Wertebereich annehmen. Zu beachten ist hierbei, dass die kostengünstigeren unternehmensinternen Speichermedien längere Zugriffszeiten aufweisen, sodass das Abrufen von unstrukturierten Daten bzw. Dokumenten mit längeren Wartezeiten verbunden ist.¹¹⁷² Bislang existierten keine wissenschaftlichen Untersuchungen dazu, wie das Cloud Computing im ILM eingesetzt werden kann, um unstrukturierte Daten wirtschaftlich zu verwalten.¹¹⁷³ Wenn es die rechtlichen Anforderungen¹¹⁷⁴ zulassen, sollten die Daten möglichst in der Cloud-Ebene verwaltet werden,¹¹⁷⁵ da das Kostensenkungspotential hier am größten ist.¹¹⁷⁶ Die Klassifizierung der Daten kann nur teilweise automatisiert werden, da die Daten vor der Klassifizierung manuell entsprechenden Dokumentenklassen zugeordnet werden müssen. Somit kann vermieden werden, dass zusammengehörige Daten, die bspw. für den Ablauf eines Geschäftsprozesses notwendig sind, durch Konzepte zur automatisierten Klassifizierung auf unterschiedliche Speicherebenen abgelegt werden und ein reibungsloser Ablauf des Geschäftsprozesses verhindert wird.¹¹⁷⁷ Hierzu müssen hinsichtlich einer praktischen Implementierung eines ILM-Konzepts Anwender und Administratoren bezüglich der Verwendung der Daten befragt werden.¹¹⁷⁸

Die Anwendung des in dieser Arbeit konstruierten Verfahrens zur Wirtschaftlichkeitsanalyse der automatisierten Klassifizierung ist ebenfalls für strukturierte Daten denkbar.¹¹⁷⁹ Ein großer Teil der strukturierten Informationen eines Unternehmens befindet sich innerhalb von Datenbanken, beispielsweise in Data-Warehouse-Systemen. Analog

¹¹⁷² Vgl. Kapitel 2.4.2

¹¹⁷³ Vgl. Sotomayor et al. / Private and Hybrid Clouds / 14-21; Jo et al. / Virtual Disk / 375-377; Hill et al. / Guide Cloud Computing / 250-251

¹¹⁷⁴ Vgl. Kapitel 3.4.2

¹¹⁷⁵ In dieser Arbeit wurde die Cloud-Ebene für die Abdeckung von Lastspitzen verwendet. Vgl. hierzu Kapitel 5.2.3.1.1. Denkbar wäre, die Ebenen „Online“ und „Nearline“ durch die Cloud-Ebene zu ersetzen. Eine Reduktion der Anzahl der Speicherebenen würde zusätzlich zu einer Senkung der Verlagerungsvorgänge und -kosten beitragen.

¹¹⁷⁶ Vgl. hierzu Kapitel 5.3. Cloud Computing bzw. IaaS ist zugleich für das schnelle Abrufen von Daten geeignet, da die Hardwareressourcen entsprechend der Anforderungen an die Zugriffe flexibel skalieren. Vgl. hierzu Os, Bressan / Cloud Real-Time Application / 366-367; Mahmood / Cloud Computing / 307; Deka / Big Data Analytics / 353.

¹¹⁷⁷ Vgl. Kapitel 5.2.2.1.3.1

¹¹⁷⁸ Vgl. Kapitel 5.1.3.2.3

¹¹⁷⁹ Vgl. Kosler, Matthesius, Stelzer / Klassifizierung von Informationen / 129-145

zur Vorgehensweise des in Kapitel 5.1.2 konstruierten Vorgehensmodells kann zunächst das Data Warehouse untersucht werden. Hierbei müssen die Besonderheiten eines Data Warehouse berücksichtigt werden, da die Verbindungen zwischen den Systemen sowie entwickelte Analyseberichte bzw. Abfragen¹¹⁸⁰ wichtige Informationen zu den Daten und zur Nutzung der Daten liefern. Die Klassifizierung der Daten kann ebenfalls mit Hilfe des Nutzungsgrads erfolgen. Die Klassifizierungsfunktionen müssen derart modifiziert werden, dass die von den Nutzern analysierten bzw. abgefragten Daten der Faktentabelle der Datenbank bzw. des multidimensionalen Datenwürfels untersucht werden können.¹¹⁸¹ Auf diese Weise können die Zugriffe im Zeitverlauf analysiert werden, um anschließend die Daten den entsprechenden Speicherebenen zuzuordnen.¹¹⁸² Die Analyse einzelner Datensätze einer Datenbank hinsichtlich des Nutzungsgrads ist mit einem größeren Aufwand verbunden, als die Analyse der Metadaten unstrukturierter Daten bzw. Dokumente.¹¹⁸³ Eine wichtige Voraussetzung hierfür ist, dass die statistischen Informationen über die Ausführung von Analyseberichten aufgezeichnet werden.¹¹⁸⁴ Ähnlich wie bei der automatisierten Klassifizierung unstrukturierter Daten müssen die entsprechenden Metadaten vorhanden sein oder protokolliert werden können. Zur Wirtschaftlichkeitsanalyse der automatisierten Klassifizierung von strukturierten Daten kann das in Kapitel 5.2.3.3 vorgestellte Kostenmodell zugrunde gelegt werden. Die relevanten Kosten bezüglich eines Data Warehouse können den jeweiligen Kostenarten zugeordnet werden. Die Kostenfunktionen zur Berechnung der Speicherplatz-, Administrations- und Verlagerungskosten sind den Gegebenheiten eines Data Warehouse anzupassen.

Die vorliegende Arbeit erweitert den Stand des Wissens zum ILM dahingehend, dass erstmals Aussagen zur Wirtschaftlichkeit der automatisierten Verwaltung unstrukturierter

¹¹⁸⁰ Analyseberichte bzw. Abfragen dienen der Entscheidungsunterstützung.

¹¹⁸¹ Vgl. Kosler, Matthesius, Stelzer / Klassifizierung von Informationen / 136-137

¹¹⁸² Das Data Warehouse, wie bspw. das SAP Business Information Warehouse, stellt hierfür Schnittstellen für Nearline-Storage-Systeme und Archivsysteme zur Verfügung. Vgl. hierzu Steinbinder, Bedau, Löw / Datenbanktechnologie für das ILM / 101-114.

¹¹⁸³ Die Datensätze und relevanten Datenbankobjekte unterliegen starken Abhängigkeiten mit wiederum anderen Datenbankobjekten aufgrund der mehrdimensionalen Modellierung in einem Data Warehouse. Vgl. hierzu Chaudhuri, Weikum / Database System Architecture / 1-10; Kaiser, Smolnik, Riempp / ILM-Framework / 484.

¹¹⁸⁴ Im SAP Business Warehouse wird hierzu der sogenannte Business Intelligence Content zur Verfügung gestellt.

ter Daten im ILM getroffen werden können. Um die wesentlichen Teilaufgaben des ILM sowie die Vorgehensweise bei der Klassifizierung von Informationen zu beschreiben, wurde ein ILM-Vorgehensmodell konstruiert. Es wurde untersucht, inwiefern die automatisierte Klassifizierung unstrukturierter Daten im ILM zu einer Kostensenkung in Unternehmen beitragen kann und welche Kosten- und Nutzenaspekte dabei von Relevanz sind. Die Wirtschaftlichkeit im ILM wurde für unterschiedliche Lebenszyklen und Datenmengen gemessen und ausgewertet.¹¹⁸⁵ Hierzu wurde ein Kostenmodell entwickelt, das die Anschaffungskosten, Betriebskosten und sonstigen Kosten für das ILM enthält. Das Konzept des ILM wurde in dieser Arbeit erweitert, indem nicht nur unternehmensinterne Speichermedien betrachtet wurden, sondern zusätzlich die Verwaltung von Informationen im Cloud Computing. Es wurde aufgezeigt, welche Auswirkungen das Cloud Computing auf die Wirtschaftlichkeit bezüglich der Verwaltung von Informationen im ILM hat. Weiterhin wurde untersucht, inwieweit relevante rechtliche Vorschriften hinsichtlich der Verwaltung von Informationen für das ILM von Relevanz sind.

6.2 Ausblick

ILM hat noch nicht den Stand erreicht, sich in der Praxis durchzusetzen,¹¹⁸⁶ da umfangreiche und vollständige Implementierungen bislang fehlen.¹¹⁸⁷ Im Fokus zukünftiger Forschung und Entwicklung steht die Weiterentwicklung von ILM-Software.¹¹⁸⁸ Eine Software, die automatisiert Vorschläge bezüglich der jeweils wirtschaftlich sinnvollsten Speicherebene liefert, auf der die Daten gespeichert werden sollen, wäre für die zukünftige Entwicklung und Umsetzung von ILM-Konzepten von Vorteil.¹¹⁸⁹ Die SNIA liefert hierzu mit der Entwicklung des eXtensible Access Method-Standards (XAM) einen ersten Ansatzpunkt.¹¹⁹⁰ Durch die Einbeziehung verschiedener Metadaten wie Größe, Alter und Zugriffshäufigkeit kann die Klassifizierung der Daten unterstützt werden, um

¹¹⁸⁵ Vgl. hierzu die Forschungsfragen der Arbeit Kapitel 1.3.1

¹¹⁸⁶ Vgl. Alatorre et al. / Intelligent ILM / 9-18

¹¹⁸⁷ Vgl. Möller / Lifecycle models / 67-88; Du et al. / Optimal Storage / 731-735

¹¹⁸⁸ Vgl. Alatorre et al. / Intelligent ILM / 9-18; Jeong et al. / Effective Storage / 30-46

¹¹⁸⁹ Dies würde den Aufwand der Klassifizierung der Daten und die Administrations- und Betriebskosten im ILM reduzieren.

¹¹⁹⁰ Vgl. Thome, Sollbach / Modelle des ILM / 188-189; Sosinsky / Cloud Computing Bible / 333; Giarretta / Digital Preservation / 307

bspw. Vorschläge für das jeweils sinnvollste Speichermedium zu erzeugen.¹¹⁹¹ Denkbar wäre in diesem Zusammenhang die Anwendung von Methoden zur Inhaltserschließung¹¹⁹² von unstrukturierten Daten.¹¹⁹³

Die Berücksichtigung von rechtlichen Anforderungen durch Funktionen zur automatisierten Klassifizierung ist ein offenes Forschungsthema.¹¹⁹⁴ Ansätze hierfür liefern Mont und Park.¹¹⁹⁵ Die Autoren definieren Funktionen und Regeln, die vorgeben, wie die Informationen aufgrund rechtlicher Anforderungen behandelt werden sollen. Die in den Regeln definierten Aktionen beschreiben, ob vertrauliche Informationen verlagert, kopiert oder gelöscht werden sollen. Die automatisierte Erstellung von Regeln zu rechtlichen Anforderungen, bspw. auf der Grundlage rechtlicher Vorschriften oder Gesetzesänderungen, ist nicht möglich. Diese Aufgabe obliegt den Anwendern und Administratoren.¹¹⁹⁶ Die Ansätze tragen zumindest teilweise zu einer Berücksichtigung der rechtlichen Anforderungen bei. Deshalb empfiehlt sich die Aufnahme dieser Ansätze in ein zukünftig zu entwickelndes Konzept zur automatisierten Klassifizierung von Informationen, welches zugleich die Anforderungen aus Kapitel 5.1.3.2 erfüllt.

Um die Kostensenkungspotentiale des Cloud Computings im ILM zu nutzen, empfiehlt sich die Speicherung der Informationen in einer Public Cloud.¹¹⁹⁷ Schwerpunkt zukünftiger Forschung ist die Schaffung von Sicherheitsstandards zur Gewährleistung der ordnungsgemäßen Übertragung, Aufbewahrung und Wiedergabe der Informationen im Cloud Computing.¹¹⁹⁸ Die Informationen sind vor Beeinträchtigungen, wie menschli-

¹¹⁹¹ Der XAM-Standard allein sorgt nicht für die automatisierte Klassifizierung von Informationen und müsste entsprechend um die unternehmensspezifischen Eigenschaften und Anforderungen erweitert werden, wie bspw. Betriebskosten, Administrationskosten, Zugriffsintervalle oder verwendete Dokumentenklassen. Allerdings sorgt der Standard dafür, dass Unternehmen nicht von einzelnen ILM-Anbietern abhängig sind, die bspw. zukünftig ILM-Software mit proprietären Formaten und Schnittstellen anbieten. Vgl. hierzu Poulton / Data Storage Networking / 279.

¹¹⁹² Vgl. Fugmann / Inhaltserschließung / 101-106

¹¹⁹³ Vgl. hierzu Cheptsov et al. / Scalable Cloud Platform / 8-12. Damit könnte die Vorklassifizierung der Dokumente unterstützt werden. Vgl. hierzu Kapitel 5.2.2.1.3.1.

¹¹⁹⁴ Vgl. Hadjiantonis, Pavlou / Self-Management / 257-258; Krishnan / Big Data / 113-114; Shen et al. / Data Protection in the Cloud / 137-139; Mont et al. / Privacy Policy Enforcement / 138

¹¹⁹⁵ Vgl. Mont / Identity Lifecycle Management / 397-426; Park / Privacy Policy Grid / 91-100; Park, Kim / Policy energy management / 653-664

¹¹⁹⁶ Vgl. Hynes, Reynolds, Hauswirth / Context Lifecycle / 56-57

¹¹⁹⁷ Vgl. Kapitel 5.3.7

¹¹⁹⁸ Vgl. Demchenko et al. / Security Cloud Infrastructure / 183; Urbanski / Cloud-Outcomes / 64

ches und technisches Versagen, Fehler bei der Informationsübertragung, Spionage¹¹⁹⁹ oder höhere Gewalt, zu schützen.¹²⁰⁰ Zur Sicherstellung der Leistung und angemessenen Nutzung ihrer IT-Infrastruktur halten Anbieter von Public-Cloud-Diensten die gespeicherten Informationen ihrer Kunden redundant oder transferieren diese oft in unterschiedliche Rechenzentren.¹²⁰¹ Dabei werden die Informationen oft auf Systemen von Subunternehmern des Anbieters gespeichert und es werden Landesgrenzen überschritten.¹²⁰² Hierin liegt eine wesentliche Problematik des Cloud Computings hinsichtlich der Transparenz der Datenverarbeitung, weshalb eine Speicherung von geschäftskritischen oder vertraulichen Informationen derzeit nicht angemessen ist.¹²⁰³ Der Cloud-Anbieter kann nicht immer eine durchgängige Verschlüsselung der Informationen garantieren.¹²⁰⁴ Weiterhin geben die meisten Cloud-Anbieter keine Zusicherungen darüber, in welchem Land die Informationen gespeichert werden und welche Maßnahmen ergriffen werden, um die Vertraulichkeit der Informationen zu gewährleisten.¹²⁰⁵

¹¹⁹⁹ Vgl. hierzu die Ausführungen zum Schutz vor unbefugtem Zugriff auf die Informationen in einer Public Cloud Kapitel 3.4.2.

¹²⁰⁰ Vgl. BSI / Sicherheitshandbuch / ; Vgl. Kapitel 3.4.1

¹²⁰¹ Vgl. Bauer, Adams / Service Quality Cloud / 93-95

¹²⁰² Vgl. Buyya, Broberg, Goscinski / Cloud Principles / 603-605

¹²⁰³ Vgl. Bo / Security Problems Cloud / 2497; Hill et al. / Guide Cloud Computing / 260; Onwubiko / Security Issues Cloud / 285

¹²⁰⁴ Vgl. Kriesi / Rechtliche Aspekte von SaaS / 40-41; Linthicum / Cloud Contracts /

¹²⁰⁵ Vgl. Kriesi / Rechtliche Aspekte von SaaS / 40-41; Gantz, Reinsel / Extracting Value / 9

Literaturverzeichnis

Abd-El-Malek et al. / Cluster-based storage /

Michael Abd-El-Malek, William Courtright, Chuck Cranor, Gregory Ganger, James Hendricks, Andrew Klosterman, Michael Mesnier, Manish Prasad, Brandon Salmon, Rajan Sambasivan, Shafeeq Sinnamohideen, John Strunk, Eno Thereska, Matthew Wachs, Jay Wylie: Ursa Minor: versatile cluster-based storage. In: Proceedings of the 4th USENIX Conference on File and Storage Technologies. San Francisco 2005, S. 59-72.

Abd-El-Malek et al. / self-storage /

Michael Abd-El-Malek, William Courtright, Chuck Cranor, Gregory Ganger, James Hendricks, Andrew Klosterman, Michael Mesnier, Manish Prasad, Brandon Salmon, Rajan Sambasivan, Shafeeq Sinnamohideen, John Strunk, Eno Thereska, Matthew Wachs, Jay Wylie: Early experiences on the journey towards self-* storage. In: Lomet (Hrsg.): Data Engineering Bulletin, Band 29, Nr. 4, Atlanta 2006, S. 55-62.

Abiteboul / Semi-Structured Data /

Serge Abiteboul: Querying Semi-Structured Data. In: Proceedings of the 6th International Conference on Database Theory. Delphi 1997, S. 1-18.

Agrawal et al. / Study File-System Metadata /

Nitin Agrawal, William Bolosky, John Douceur, Jacob Lorch: A Five-Year Study of File-System Metadata. In: Proceedings of the 5th USENIX Conference on File and Storage Technologies. New York 2007, S. 31-45.

Aichele / Smart Energy /

Christian Aichele: Smart Energy. Von der reaktiven Kundenverwaltung zum proaktiven Kundenmanagement. Wiesbaden 2012.

AIS / MIS Journal Ranking /

Association for Information Systems: MIS Journal Rankings.
<http://aisnet.org/?JournalRankings>, Atlanta 2014,
Abruf: 2014-01-14.

Al-Fedaghi / ILM /

Sabah Al-Fedaghi: On Information Lifecycle Management. In: Proceedings of the 2008 IEEE Asia-Pacific Services Computing Conference. Yilan 2008, S. 335-342.

Alatorre et al. / Intelligent ILM /

Gabriel Alatorre, Aameek Singh, Nagapramod Mandagere, Eric Butler, Sandeep Gopisetty, Yang Song: Intelligent Information Lifecycle Management in Virtualized Storage Environments. In: Proceedings of the 2014 Annual SRII Global Conference. San Jose 2014, S. 9-18.

Alisch / Wirtschaftslexikon /

Katrin Alisch (Hrsg.): Gabler-Wirtschaftslexikon, Band 3. 16. Auflage, Wiesbaden 2004.

Allaire et al. / Costs Data Migrations /

Patrick Allaire, Justin Augat, Joe Jose, David Merrill: Reduce Costs and Risks for Data Migrations. Data Migration Best Practices and Nondisruptive Migration Service Capability for Enterprise Storage. Santa Clara 2012.

Allen / waste storage dollars /

Nick Allen: Don't waste your storage dollars: what you need to know. Research Note, Gartner Group Inc., Stamford 2001.

Allman, Laurito, Loh / Simulation Modeling in Excel /

Keith Allman, Josh Laurito, Michael Loh: Financial Simulation Modeling in Excel + Website. A Step-by-Step Guide. 1. Auflage, Hoboken 2011.

Allweyer / Geschäftsprozessmanagement /

Thomas Allweyer: Geschäftsprozessmanagement. Strategie, Entwurf, Implementierung, Controlling. Bochum 2008.

Amazon / AWS Import und Export /

Amazon: AWS Import/Export.
<http://aws.amazon.com/de/importexport/>; Seattle 2014,
Abruf: 2014-09-26.

Amazon / Calculator /

Amazon: Simple Monthly Calculator.
<http://calculator.s3.amazonaws.com/index.html>; Seattle 2014,
Abruf: 2014-09-26.

Amazon / Import-Export Preise /

Amazon: AWS Import/Export - Preise.
<http://aws.amazon.com/de/importexport/pricing/>; Seattle 2014,
Abruf: 2014-09-26.

Amazon / S3 Price Reduction /

Amazon: Amazon Web Services Blog: Amazon S3 Price Reductions.
<http://aws.amazon.com/de/blogs/aws/aws-price-reduction-42-ec2-s3-rds-elasticache-and-elastic-mapreduce/>; Seattle 2014,
Abruf: 2014-10-07.

Amazon S3 / S3-Preise /

Amazon Simple Storage Service: Amazon S3 - Preise.
<http://aws.amazon.com/de/s3/pricing/>; Seattle 2014,
Abruf: 2014-09-24.

Amberg, Wiener / IT-Offshoring /

Michael Amberg, Martin Wiener: IT-Offshoring. Management internationaler IT-Outsourcing-Projekte. Heidelberg 2006.

Andree / Wirtschaftlichkeitsanalyse /

Ulrich Andree: Wirtschaftlichkeitsanalyse öffentlicher Investitionsprojekte. Investitionen sicher und zuverlässig planen. Freiburg 2011.

Antonovich / Office User's Guide /

Michael Antonovich: Office and SharePoint 2010 User's Guide. Integrating SharePoint with Excel, Outlook, Access and Word. 1. Auflage, New York, 2010.

Antweiler / Wirtschaftlichkeitsanalyse /

Johannes Antweiler: Wirtschaftlichkeitsanalyse von Informations- und Kommunikationssystemen (IKS) - Wirtschaftlichkeitsprofile als Entscheidungsgrundlage. Köln 1995.

Archer / Method for designers /

Bruce Archer: Systematic method for designers. In: Nigel Cross (Hrsg.): Developments in Design Methodology. London 1984.

Arimura / Patterns from Semi-structured Data /

Hiroki Arimura: Efficient Algorithms for Mining Frequent and Closed Patterns from Semi-structured Data. In: Takashi Washio, Einoshin Suzuki, Kai Ming Ting, Akihiro Inokuchi (Hrsg.): Proceedings of the 12th Pacific-Asia Conference on Knowledge Discovery and Data Mining. Osaka 2008, S. 2-13.

Armbrust et al. / Cloud Computing /

Michael Armbrust, Armando Fox, Rean Griffith, Anthony Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, Matei Zaharia: A Berkeley View of Cloud Computing. Technical Report, Nr. 28, Universität Berkeley 2009.

AWV / Aufbewahrungsverfahren /

Arbeitskreis für wirtschaftliche Verwaltung e.V.: Gesetzliche Anforderungen an Aufbewahrungsverfahren und Speichermedien. Eschborn 1989.

AWV / Leitlinien /

Arbeitskreis für wirtschaftliche Verwaltung e.V.: Organisatorische Leitlinien für den Einsatz von optischen Speichermedien. Eschborn 1993.

AWV / Mikrofilm /

Arbeitskreis für wirtschaftliche Verwaltung e.V.: Der Mikrofilm als Aufzeichnungsmedium in Revision und Betriebsprüfung. Berlin 1983.

AWV / Speicherplattensysteme /

Arbeitskreis für wirtschaftliche Verwaltung e.V.: Rechtliche Rahmenbedingungen für den Einsatz optischer Speicherplattensysteme. Eschborn 1992.

AWV / Verfahrensdokumentation /

Arbeitskreis für wirtschaftliche Verwaltung e.V.: Zulässigkeit, Ordnungsmäßigkeit und Verfahrensdokumentation der Schriftgutverfilmung nach Handels- und Steuerrecht. Eschborn 1987.

Backhaus et al. / Analysemethoden /

Klaus Backhaus, Bernd Erichson, Wulff Plinke, Rolf Weiber: Multivariate Analysemethoden. Eine anwendungsorientierte Einführung. 11. Auflage, Berlin - Heidelberg - New York 2005.

Badach, Rieger / Netzwerk-Projekte /

Anatol Badach, Sebastian Rieger: Netzwerk-Projekte. Planung, Realisierung, Dokumentation und Sicherheit von Netzwerken. München 2013.

Badam, Kumari, Srinivasan / ILM healthcare /

Nitin Badam, Aradhna Kumari, Jagannathan Srinivasan: Information Lifecycle Management in evolving healthcare databases. In: COMAD '13 Proceedings of the 19th International Conference on the Management of Data, Mumbai 2013, S. 79-90.

Badertscher, Gubelmann, Scheuring / Informationssysteme /

Kurt Badertscher, Josef Gubelmann, Johannes Scheuring: Wirtschaftsinformatik Grundlagen: Informations- und Kommunikationssysteme gestalten. Zürich 2006.

Baier / Controlling /

Peter Baier: Praxishandbuch Controlling. Controlling-Instrumente, Unternehmensplanung und Reporting. 2. Auflage, München 2008.

Baker / Adobe Acrobat /

Donna Baker: Adobe Acrobat 5. 1. Auflage, Berlin - Heidelberg - New York, 2002.

Ballou, Smith, Wang / Quality Estimation Database /

Donald Ballou; Indu Shobha Smith, Richard Wang: Sample-Based Quality Estimation of Query Results in Relational Database Environments. In: IEEE Transactions on Knowledge and Data Engineering, Band 18, Nr. 5, 2006, S. 639-650.

Balzert /Software-Technik/

Helmut Balzert: Lehrbuch der Software-Technik. 2. Auflage, Heidelberg - Berlin 2001.

Banks / Simulation /

Jerry Banks: Handbook of Simulation. Principles, Methodology, Advances, Applications and Practice. 1. Auflage, New York 1998.

Bauer, Adams / Service Quality Cloud /

Eric Bauer, Randee Adams: Service Quality of Cloud-Based Applications. 1. Auflage, Hoboken 2014.

Baun et al. / Cloud Computing /

Christian Baun, Marcel Kunze, Jens Nimis, Stefan Tai: Cloud Computing. Web-Based Dynamic IT Services. 2. Auflage, Berlin - Heidelberg 2011.

Baun, Kunze, Ludwig / Servervirtualisierung /

Christian Baun, Marcel Kunze, Thomas Ludwig: Servervirtualisierung. In: Informatik-Spektrum, Band 32, Nr. 3, 2009, S. 197-205.

Bechmann / Nutzwertanalyse /

Arnim Bechmann: Nutzwertanalyse, Bewertungstheorie und Planung. Stuttgart 1978.

Beck et al. / Requirements Engineering /

Martina Beck, Christian Loos, Volker Maiborn, Dominik Neff: Requirements Engineering in frühen Phasen. In: Michael Lang, Sebastian Kammerer, Michael Amberg (Hrsg.): Perfektes IT-Projektmanagement. Best Practices für Ihren Projekterfolg. Düsseldorf 2012, S. 19-42.

Becker / Finanzwirtschaft /

Hans Becker: Investition und Finanzierung. Grundlagen der betrieblichen Finanzwirtschaft. 4. Auflage, Wiesbaden 2010.

Becker / Wirtschaftswissen /

Hans Becker: Investieren und Finanzieren. In: Lutz Irgel (Hrsg.): Gablers Wirtschaftswissen für Praktiker. Zuverlässig Orientierung in allen kaufmännischen Fragen. 5. Auflage, Wiesbaden 2004.

Becker et al. / Referenzmodellierung /

Jörg Becker, Patrick Delfmann, Ralf Knackstedt, Dominik Kuropka: Konfigurative Referenzmodellierung. In: Jörg Becker, Ralf Knackstedt (Hrsg.): Wissensmanagement mit Referenzmodellen. Konzepte für die Anwendungssystem- und Organisationsgestaltung. Heidelberg 2002, S. 25-144.

Becker et al. / Referenzvorgehensmodell /

Jörg Becker, Lars Algermissen, Patrick Delfmann, Björn Niehaves: Prozessorientierte Reorganisation in öffentlichen Verwaltungen - Erfahrungen bei der Anwendung eines Referenzvorgehensmodells. In: Jörg Becker, Patrick Delfmann (Hrsg.): Referenzmodellierung. Grundlagen, Techniken und domänenbezogene Anwendungen. Berlin - Heidelberg 2004, S. 151-177.

Becker, Meise / Strategie und Ordnungsrahmen /

Jörg Becker, Volker Meise: Strategie und Ordnungsrahmen. In: Jörg Becker, Martin Kugler, Michael Rosemann (Hrsg.): Prozessmanagement. Ein Leitfaden zur prozessorientierten Organisationsgestaltung. 7. Auflage, Berlin - Heidelberg 2012, S. 113-164.

Becker, Pfeiffer / Forschung in der WI /

Jörg Becker, Daniel Pfeiffer: Beziehungen zwischen behavioristischer und konstruktionsorientierter Forschung in der Wirtschaftsinformatik. In: Stephan Zelewski, Naciye Akca (Hrsg.): Fortschritt in den Wirtschaftswissenschaften - Wissenschaftstheoretische Grundlagen und exemplarische Anwendungen. Wiesbaden 2006, S. 1-17.

Becker, Pfeiffer / Konzeptuelle Modellierung /

Jörg Becker, Daniel Pfeiffer: Konzeptuelle Modellierung - ein wissenschaftstheoretischer Forschungsleitfaden. In: Franz Lehner, Holger Nösekabel, Peter Kleinschmidt (Hrsg.): Multikonferenz Wirtschaftsinformatik 2006. Berlin 2006, S. 3-19.

Becker, Probandt, Vering / Grundsätze ordnungsgemäßer Modellierung /

Jörg Becker, Wolfgang Probandt, Oliver Vering: Grundsätze ordnungsgemäßer Modellierung. Konzeption und Praxisbeispiel für ein effizientes Prozessmanagement. Berlin - Heidelberg 2012.

Becker, Schütte / Handelsinformationssysteme /

Jörg Becker, Reinhard Schütte: Handelsinformationssysteme. 2. Auflage, Frankfurt am Main 2004.

Beinhauer, Herr, Schmidt / SOA /

Wolfgang Beinhauer, Michael Herr, Achim Schmidt: Warum SOA? Eine Einführung. In: Wolfgang Beinhauer, Michael Herr, Achim Schmidt (Hrsg.): SOA für agile Unternehmen. Serviceorientierte Architekturen verstehen, einführen und nutzen. Düsseldorf 2008, S. 19-40.

Bennett, Bauer, Kinchlea / NFS environments /

Michael Bennett, Michael Bauer, David Kinchlea: Characteristics of files in NFS environments. In: ACM SIGSMALL/PC Notes, Band 18, Nr. 3-4, 1992, S. 18-25.

Bergmann / Speichermedien /

M. Bergmann: Sind die Hürden überwunden? Vom Mikrofilm zu digitalen Speichermedien. In: Infodoc. Nr.2, 1992, S. 21-24.

Bernard / ILM Security /

Ray Bernard: Information Lifecycle Security Risk Assessment: A tool for closing security gaps. In: Journal of Computer and Security, Band 26, Nr. 1, 2007, S. 26-30.

Betzler / Virtualisierung /

Boas Betzler: Virtualisierung. In: Herbert Kircher (Hrsg.): IT-Technologien, Lösungen, Innovationen. Berlin - Heidelberg - New York 2007, S. 100-110.

Bhagwan et al. / Management of Data Storage /

Ranjita Bhagwan, Fred Douglass, Kirsten Hildrum, Jeffrey Kephart, William Walsh: Time-varying Management of Data Storage. In: Candea, Oppenheimer (Hrsg.): First Workshop on Hot Topics in Systems Dependability, Yokohama, 2005.

Bieg, Kußmaul / Finanzierung /

Hartmut Bieg, Heinz Kußmaul: Finanzierung. 2. Auflage, München 2009.

Biethahn, Mucksch, Ruf / Informationsmanagement /

Jörg Biethahn, Harry Mucksch, Walter Ruf: Ganzheitliches Informationsmanagement. Band 1: Grundlagen. 6. Auflage, München 2004.

BITKOM / Leitfaden Cloud Computing /

Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V.: BITKOM-Leitfaden: Cloud Computing – Evolution in der Technik, Revolution im Business.

http://www.bitkom.org/files/documents/BITKOM-Leitfaden-CloudComputing_Web.pdf; Berlin 2009, Abruf: 2013-02-07.

BITKOM / Leitfaden ILM /

Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V.: Leitfaden zum Thema „Information Lifecycle Management“.
http://www.bitkom.org/files/documents/BITKOM_Leitfaden_ILM_Stand_21-04-2004.pdf, Berlin 2004,
Abruf: 2011-04-08.

BITKOM / Server-Virtualisierung /

Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V.: Server-Virtualisierung. Teil 2: Design, Deployment und Betrieb.
http://www.bitkom.org/files/documents/virtualisierung_nov_2009_T2.pdf, Berlin 2009,
Abruf: 2013-05-31.

Böhm, Fuchs / System-Entwicklung /

Rolf Böhm, Emmerich Fuchs: System-Entwicklung in der Wirtschaftsinformatik. 5. Auflage, Zürich 2002.

Bo / Security Problems Cloud /

Li Bo: Security Problems and Strategies of Digital Education Resource Management in Cloud Computing Environment. In: Shaozi Li, Qun Jin, Xiaohong Jiang, James Park (Hrsg.): Frontier and Future Development of Information Technology in Medicine and Education, Lecture Notes in Electrical Engineering, Band 269, Dordrecht 2014, S. 2495-2500.

Bodrow, Bergmann / Wissensbewertung /

Wladimir Bodrow, Philipp Bergmann: Wissensbewertung in Unternehmen. Bilanzieren von intellektuellem Kapital. Berlin 2003.

Bossel / Simulation komplexer Systeme /

Hartmut Bossel: Systeme, Dynamik, Simulation. Modellbildung, Analyse und Simulation komplexer Systeme. Norderstedt 2004.

Box, Muller / Random Normal Deviates /

George Box, Mervin Muller: A note on the generation of random normal deviates. In: Theodore Harris (Hrsg.): Annals of Mathematical Statistics, Band 29, Nr. 2, 1958, S. 610-611.

Bräu, Dannert / Effizienzsteigerung /

Udo Bräu, Juliane Dannert: Effizienzsteigerung im Krankenhaus - Ist der IT-Einsatz ein wesentliches Mittel zu mehr Wirtschaftlichkeit im OP? In: Helmut Schlegel (Hrsg.): Steuerung der IT im Klinikmanagement. Wiesbaden 2010, S. 253-264.

Brandt-Pook, Kollmeier / Softwareentwicklung /

Hans Brandt-Pook, Rainer Kollmeier: Softwareentwicklung kompakt und verständlich. Wie Softwaresysteme entstehen. Wiesbaden 2008.

Brockhaus Enzyklopädie / Zyklus /

Brockhaus Enzyklopädie. Band 24, 19. Auflage, Mannheim 1997.

Broy, Kuhrmann / Projektorganisation /

Manfred Broy, Marco Kuhrmann: Projektorganisation und Management im Software Engineering. Berlin - Heidelberg 2013.

Brugger / IT Business Case /

Ralph Brugger: Der IT Business Case. Kosten ermitteln und analysieren. Nutzen erkennen und quantifizieren. Wirtschaftlichkeit nachweisen und realisieren. 2. Auflage, Berlin - Heidelberg 2009.

BSI / Sicherheitsempfehlungen Cloud /

Bundesamt für Sicherheit in der Informationstechnik (BSI) (Hrsg.): Sicherheitsempfehlungen für Cloud Computing Anbieter. Mindestanforderungen in der Informationssicherheit

https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Mindestanforderungen/Eckpunktepapier-Sicherheitsempfehlungen-CloudComputing-Anbieter.pdf?__blob=publicationFile, Bonn 2012, Abruf: 2014-01-07.

BSI / Sicherheitshandbuch /

Bundesamt für Sicherheit in der Informationstechnik (BSI) (Hrsg.): IT-Sicherheitshandbuch: Handbuch für die sichere Anwendung der Informationstechnik, <https://www.bsi.bund.de/ContentBSI/Publikationen/KriterienSicherheitshandbuch/sicherheitshandbuch.html>, Bonn 1992, Abruf: 2011-04-13.

Bundesministerium der Finanzen / GDPdU /

Bundesministerium der Finanzen: Grundsätze zum Datenzugriff und zur Prüfbarkeit digitaler Unterlagen (GDPdU) vom 16. Juli 2001 (BMF-Schreiben).

Bundesministerium der Finanzen / GoBS /

Bundesministerium der Finanzen: Grundsätze ordnungsgemäßer DV-gestützter Buchführungssysteme (GoBS) vom 7. November 1995 (BStBl S. 738).

Bundesministerium der Justiz / AO /

Bundesministerium der Justiz: Abgabenordnung (AO) vom 16. März 1976, zuletzt geändert durch Art. 10 des Gesetzes vom 20. Dezember 2008.

Bundesministerium der Justiz / BDSG /

Bundesministerium der Justiz: Bundesdatenschutzgesetz (BDSG) vom 20. Dezember 1990, in der Fassung der Bekanntmachung vom 14. August 2009 (BGBl. I S. 66), zuletzt geändert durch Art. 15 Absatz 53 des Gesetzes vom 5. Februar 2009.

Bundesministerium der Justiz / GG /

Bundesministerium der Justiz: Grundgesetz der Bundesrepublik Deutschland vom 23. Mai 1949, zuletzt geändert durch Art. 1 des Gesetzes vom 11. Juli 2012.

Bundesministerium der Justiz / HGB /

Bundesministerium der Justiz: Handelsgesetzbuch (HGB) vom 10. Mai 1897, zuletzt geändert durch Art. 1 des Gesetzes vom 04. Oktober 2013.

Bundesministerium der Justiz / TKG /

Bundesministerium der Justiz: Telekommunikationsgesetz (TKG) vom 22. Juni 2004, zuletzt geändert durch Artikel 1 vom 20. Juni 2013.

Bundesministerium des Inneren / WiBe /

Bundesministerium des Inneren: Empfehlung zur Durchführung von Wirtschaftlichkeitsbetrachtungen in der Bundesverwaltung, insbesondere beim Einsatz der IT. WiBe Version 4.1, Band 92, 2007.

Burmann, Freiling, Hülsmann / Ad-hoc-Krisen /

Christoph Burmann, Jörg Freiling, Michael Hülsmann: Management von Ad-hoc-Krisen: Grundlagen-Strategien-Erfolgsfaktoren. Wiesbaden 2005.

Buyya, Broberg, Goscinski / Cloud Principles /

Rajkumar Buyya, James Broberg, Andrzej Goscinski: Cloud Computing. Principles and Paradigms. 1. Auflage, Hoboken 2011.

Buyya, Vecchiola, Selvi / Cloud Computing /

Rajkumar Buyya, Christian Vecchiola, Thamarai Selvi: Mastering Cloud Computing. New Delhi 2013.

Cao et al. / Grid Resource /

Lei Cao, Minglu Li, Henry Rong, Joshua Huang: An Ontology-Based Model for Grid Resource Publication and Discovery. In: Hai Jin, Yi Pan, Nong Xiao, Jianhua Sun (Hrsg.): Grid and Cooperative Computing. Third International Conference Wuhan, China. Berlin - Heidelberg - New York 2004, S. 448-455.

Cao, Li, Xia / Cloud Computing Architecture /

Bu-Qing Cao, Bing Li, Qi-Ming Xia: A Service-Oriented Qos-Assured and Multi-Agent Cloud Computing Architecture. In: Martin Gilje Jaatun, Gansen Zhao, Chunming Rong (Hrsg.): Cloud Computing. Proceedings of the First International Conference, CloudCom 2009. Beijing 2009, S. 644-649.

Carstensen / Investitionsrechnung /

Peter Carstensen: Investitionsrechnung kompakt. Eine Anwendungsorientierte Einführung. Wiesbaden 2008.

Chandra, Gehani, Yu / Automated Storage /

Surendar Chandra, Ashish Gehani, Xuwen Yu: Automated Storage Reclamation Using Temporal Importance Annotations. In: 27th International Conference on Distributed Computing Systems, Toronto 2007, S. 12.

Chang, Abu-Amara, Sanford / Cloud Services /

William Chang, Hosame Abu-Amara, Jessica Sanford: Transforming Enterprise Cloud Services. 1. Auflage, Heidelberg - London - New York 2010.

Chang, Wills, De Roure / Cloud Business Models /

Victor Chang, Gary Wills, David De Roure: Cloud Business Models and Sustainability: Impacts for businesses and e-Research. In: UK e-Science All Hands Meeting 2010, Software Sustainability Workshop, Cardiff 2010, S. 13-16.

Charles, Schalk, Thiel / Kostenmodelle /

Oliver Charles, Markus Schalk, Steffen Thiel: Kostenmodelle für Softwareproduktlinien. In: Informatik-Spektrum, Band 34, Nr. 4, 2011, S. 377-390.

Charnes / Financial Modeling /

John Charnes: Financial Modeling with Oracle Crystal Ball and Excel. 2. Auflage, Hoboken 2012.

Chaudhuri, Weikum / Database System Architecture /

Surajit Chaudhuri, Gerhard Weikum: Rethinking database system architecture: towards a selftuning RISC-style database system. In: International Conference on Very Large Databases, Morgan Kaufmann-Verlag, 2000, S. 1-10.

Chee, Franklin / Cloud Strategies /

Brian Chee, Curtis Franklin: Cloud Computing. Technologies and Strategies of the Ubiquitous Data Center. Boca Raton 2010.

Chen / Information valuation /

Ying Chen: Information valuation for Information Lifecycle Management. In: Proceedings of the Second International Conference on Autonomic Computing, 2005, S. 135-146.

Chen / Storage Hierarchies /

Peter Chen: Optimal file allocation in multi-level storage hierarchies. In: Proceedings of National Computer Conference and Exposition. AFIPS Conference Proceedings, Band 42, 1973, S. 277-282.

Cheptsov et al. / Scalable Cloud Platform /

Alexey Cheptsov, Axel Tenschert, Paul Schmidt, Birte Glimm, Mauricio Matthesius, Thorsten Liebig: Introducing a New Scalable Data-as-a-Service Cloud Platform for Enriching Traditional Text Mining Techniques by Integrating Ontology Modelling and Natural Language Processing. In: Proceedings of the International Workshop on Big Web Data (BigWebData 2013), Lecture Notes in Computer Science, Nanjing 2013, S. 8-12.

Clark / Storage Networks /

Tom Clark: Designing Storage Area Networks: A Practical Reference for Implementing Fibre Channel and IP SANs. 2. Auflage, Boston 2003.

Cloud Climate/ Watching the Clouds /

Cloud Climate: Watching the Clouds. Cloud Hosting Performance.
<http://cloudclimate.com>, o. O. 2014,
Abruf: 2014-01-07.

Connolly, Begg, Strachan / Database Systems /

Thomas Connolly, Carolyn Begg, Anne Strachan: Database Systems - A Practical Approach to Design, Implementation and Management, 2. Auflage, Harlow 1998.

Cook / Failing Project /

Rick Cook: How to spot a failing project. In: CIO Magazine. Framingham 2007.
http://www.luc.edu/media/lucedu/pmo/pdfs/additionalreading/How_to_spot_a_failing_project.pdf, Framingham 2007,
Abruf: 2014-06-06.

Cusumano / SaaS and Cloud Computing /

Michael Cusumano: Will SaaS and Cloud Computing become a New Industry Platform? In: Alexander Benlian, Thomas Hess, Peter Buxmann (Hrsg.): Software-as-a-Service. Anbieterstrategien, Kundenbedürfnisse und Wertschöpfungsstrukturen, 1. Auflage, Wiesbaden 2010, S. 3-14.

Dannehl / Speicherbedarf /

Siegfried Dannehl: Storage-Optimization-Service analysiert den Speicherbedarf von KMUs. In: Storage-Magazin. Sonderdruck 2005, S. 8.

Darcey, Conder / Android Application Programming /

Lauren Darcey, Shane Conder: Learning Android Application Programming for the Kindle Fire. A Hands-On Guide to Building Your First Android Application. New Jersey 2012.

Dauen / Aufbewahrungspflichten /

Sabine Dauen: Aufbewahrungspflichten. München 2004.

Daume / Finanzmathematik /

Peggy Daume: Finanzmathematik im Unterricht. Aktien und Optionen. Mathematische und didaktische Grundlagen mit Unterrichtsmaterialien. Wiesbaden 2009.

Deimel, Heupel, Wiltinger / Controlling /

Klaus Deimel, Thomas Heupel, Kai Wiltinger: Controlling. München 2013.

Deka / Big Data Analytics /

Ganesh Chandra Deka: Significance of In-Memory Computing for Real-Time Big Data Analytics. In: Pethuru Raj, Ganesh Chandra Deka (Hrsg.): Handbook of Research on Cloud Infrastructures for Big Data Analytics, Hershey 2014, S. 352-369.

Demchenko et al. / Security Cloud Infrastructure /

Yuri Demchenko, Canh Ngo, Cees de Laat, Chunmin Rong, Thomasz Wlodarczyk, Wolfgang Ziegler: Security Infrastructure for Dynamically Provisioned Cloud Infrastructure Services. In: Siani Pearson, George Yee (Hrsg.): Privacy and Security for Cloud Computing, London 2013, S. 167-210.

Douceur, Bolosky / Study of File-System /

John Douceur, William Bolosky: A Large-Scale Study of File-System Contents. In: SIGMETRICS '99, Atlanta 1999, S. 59-70.

Douglis et al. / Delete-Optimized Storage /

Fred Douglis, John Palmer, Elizabeth Richards, David Tao, William Tetzlaff, John Tracey, Jian Yin: Position: Short Object Lifetimes Require a Delete-Optimized Storage System. In: Proceedings of the 11th workshop on ACM SIGOPS European workshop: beyond the PC, ACM Press, New York 2004.

Doukas, Ntaikos, Bardis / ILM Applications /

Nikolaus Doukas, Konstantinos Ntaikos, Nikolaus Bardis: Integrated Information Life-Cycle. Data Management and Secret Key Lifecycle Management for Military Applications. In: Proceedings of the 10th WSEAS international conference on Mathematical methods, computational techniques and intelligent systems, Bucharest 2008, S. 492-497.

Dreger / Kundenzufriedenheit /

Wolfgang Dreger: Management der Kundenzufriedenheit. Renningen-Malmsheim 1999.

Du et al. / Optimal Storage /

Anna Du, Sanjukta Das, Ram Gopal, Ram Ramesh: Optimal Management of Digital Content on Tiered Infrastructure Platforms. In: Information Systems Research, Band 25, Nr. 4, 2014, S. 730-746.

Dunkel et al./ Verteilte Anwendungen /

Jürgen Dunkel, Andreas Eberhart, Stefan Fischer, Carsten Kleiner, Arne Koschel: Systemarchitekturen für verteilte Anwendungen. Client-Server, Multi-Tier, SOA, Event Driven Architecture, P2P, Grid, Web 2.0, 1. Auflage, München 2008.

Ehrlenspiel, Kiewert, Lindemann / Kostenmanagement /

Klaus Ehrlenspiel, Alfons Kiewert, Udo Lindemann: Kostengünstig entwickeln und konstruieren. Kostenmanagement bei der integrierten Produktentwicklung. 6. Auflage, Berlin-Heidelberg-New York 2007.

Eichholz, Vilkner / Wirtschaftsmathematik /

Wolfgang Eichholz, Eberhard Vilkner: Taschenbuch der Wirtschaftsmathematik. 6. Auflage, München 2013.

Elias-Linde / Humanressourcenmanagement /

Sabine Elias-Linde: Personalknappheit und nachhaltiges Humanressourcenmanagement. Analyse, Lösungsansätze und Gestaltungsmöglichkeiten. Dissertation vorgelegt am Institut für Betriebswirtschaftslehre der Universität Bremen, Bremen 2012.

Ellard et al. / File prediction /

Daniel Ellard, Michael Mesnier, Eno Thereska, Gregory Ganger, Marco Seltzer: Attribute-Based File Prediction of File Properties. Harvard 2003.

Elliott / Information Technology /

Geoffrey Elliott: Global Business Information Technology. An Integrated Systems Approach. 1. Auflage, Essex 2004.

Elmasri, Navathe / Datenbanksysteme /

Ramez Elmasri, Shamkant Navathe: Grundlagen von Datenbanksystemen. 3. Auflage, München 2009.

Erek, Schmidt, Schilling / Green-IT /

Koray Erek, Nils-Holger Schmidt, Thomas Schilling: Green-IT bei Bayer Business Services. In: Rüdiger Zarnekow, Koray Erek (Hrsg.): Research Papers in Information Systems Management, Technische Universität Berlin, Nr. 5, 2011, S. 1-30.

Fairbanks / Software Architecture /

George Fairbanks: Just Enough Software Architecture. A Risk-Driven Approach. 3. Auflage, Boulder 2012.

Falk, Griesberger, Leist / Artifact /

Thomas Falk, Philipp Griesberger, Susanne Leist: Patterns as an Artifact for Business Process Improvement - Insights from a Case Study. In: Jan vom Brocke, Riitta Hekkala, Sudha Ram, Matti Rossi (Hrsg.): Design Science at the Intersection of Physical and Virtual Design, 8th International Conference DESRIST 2013, Helsinki 2013, S. 88-104.

Feldman, Sanger / Text Mining Handbook /

Ronen Feldman, James Sanger: The Text Mining Handbook. Advanced Approaches in Analyzing Unstructured Data. New York 2007.

Ferstl, Sinz / Wirtschaftsinformatik /

Otto Ferstl, Elmar Sinz: Grundlagen der Wirtschaftsinformatik. 5. Auflage, München 2006.

Fettke / State of the art /

Peter Fettke: State-of-the-art des State-of-the-art. Eine Untersuchung der Forschungsmethode „Review“ innerhalb der Wirtschaftsinformatik. In: Wirtschaftsinformatik, Jahrgang 48, Nr. 4, 2006, S. 257-266.

Fettke, Loos / Referenzmodelle /

Peter Fettke, Peter Loos: Methoden zur Wiederverwendung von Referenzmodellen - Übersicht und Taxonomie. In: Jörg Becker, Ralf Knackstedt (Hrsg.): Referenzmodellierung 2002. Methoden-Modelle-Erfahrungen. Arbeitsberichte des Instituts für Wirtschaftsinformatik der Westfälischen Wilhelms-Universität Münster, Arbeitsbericht Nr. 90, Münster 2002, S. 9-34.

Fey / Grid-Computing /

Dietmar Fey: Grid-Computing: Ein Basistechnologie für Computational Science. Berlin - Heidelberg 2010.

Feyhl / Controlling von Softwareprojekten /

Achim Feyhl: Management und Controlling von Softwareprojekten. Software wirtschaftlich auswählen, entwickeln, einsetzen und nutzen. 2. Auflage, Wiesbaden 2004.

Filz / IT-Kompetenz /

Daniela Filz: Soziale IT-Kompetenz - Mythos oder Wahrheit? In: Sebastian Kammerer, Michael Amberg, Michael Lang (Hrsg.): Führung im IT-Projekt. Fachliche und soziale Kompetenzen für den Projekterfolg. Düsseldorf 2012, S. 41-66.

Finanzen.net / Eurokurs /

Finanzen.net: Eurokurs, Dollar - Euro.
<http://www.finanzen.net/devisen/eurokurs>; Karlsruhe 2014,
Abruf: 2014-10-02.

Fink / Strategische Unternehmensberatung /

Dietmar Fink: Strategische Unternehmensberatung. München 2009.

Fink, Schneidereit, Voß / Wirtschaftsinformatik /

Andreas Fink, Gabriele Schneidereit, Stefan Voß: Grundlagen der Wirtschaftsinformatik. 2. Auflage, Heidelberg 2005.

Forrest, Kaplan, Kindler / Data centers /

William Forrest, James Kaplan, Noah Kindler: Data centers: How to cut carbon emissions and costs. In: The McKinsey Quarterly, 2008, S. 1-10.

Forrest, Kaplan, Kindler / Energy Efficiency /

William Forrest, James Kaplan, Noah Kindler: Revolutionizing Data Center Energy Efficiency. In: The McKinsey Quarterly, 2008, S. 1-13.

Foscht, Angerer, Swoboda / Mixed Methods /

Thomas Foscht, Thomas Angerer, Bernhard Swoboda: Mixed Methods. Systematisierung von Untersuchungsdesigns. In: Renate Buber, Hartmut Holzmüller (Hrsg.): Qualitative Marktforschung. Konzepte-Methoden-Analysen. 2. Auflage, Wiesbaden 2009, S. 247-260.

Foster, Kesselmann, Tuecke / Anatomy of the Grid /

Ian Foster, Carl Kesselmann, Steven Tuecke: The Anatomy of the Grid: Enabling Scalable Virtual Organizations. In: International Journal of High Performance Computing Applications, Band 15, Nr. 3, 2001, S. 200-222.

Frank / Formale Sprachen /

Ulrich Frank: Zur Verwendung formaler Sprachen in der Wirtschaftsinformatik: Notwendiges Merkmal eines wissenschaftlichen Anspruchs oder Ausdruck eines übertriebenen Szientismus? In: Jörg Becker, Wolfgang König, Reinhard Schütte, Oliver Wendt, Stephan Zelewski (Hrsg.): Wirtschaftsinformatik und Wissenschaftstheorie. Bestandsaufnahme und Perspektiven. Wiesbaden 1999, S. 127-160.

Frank / Research Methods in IS /

Ulrich Frank: Towards a Pluralistic Conception of Research Methods in Information Systems Research. ICB Research Report, Nummer 7, Universität Duisburg-Essen 2006.

Frenzel, Frenzel /

Carroll Frenzel, John Frenzel: Management of Information Technology. 4. Auflage, Boston 2004.

Friedrich / TCO /

Pavel Friedrich: Allgemeine TCO-Betrachtung zum aktuellen IBM-Mainframe System z10. In: e-Journal of Practical Business Research, Band 6, 2008, S. 1-19.

Fritz / Ökonomischer Nutzen /

Sigrun Fritz: Ökonomischer Nutzen „weicher“ Kennzahlen. (Geld-) Wert von Arbeitszufriedenheit und Gesundheit. 2. Auflage, Zürich 2006.

Fugmann / Inhaltserschließung /

Robert Fugmann: Inhaltserschließung durch Indexieren: Prinzipien und Praxis. Informationswissenschaft der DGD, Band 3, Frankfurt a. M. 1999.

Fujimoto, Perumalla, Riley / Network Simulation /

Richard Fujimoto, Kalyan Perumalla, George Riley: Network Simulation. Berkeley 2007.

Gadatsch, Frick / SAP-BI /

Andreas Gadatsch, Detlev Frick: SAP-gestütztes Rechnungswesen: methodische Grundlagen und Fallbeispiele mit mySAP ERP und SAP-BI. 2. Auflage, Wiesbaden 2005.

Gadatsch, Mayer / IT-Controlling /

Andreas Gadatsch, Elmar Mayer: Masterkurs IT-Controlling. 5. Auflage, Wiesbaden 2014.

Gantz et al. / Digital Universe /

John Gantz, David Reinsel, Christopher Chute, Wolfgang Schlichting, John McArthur, Stephen Minton, Irida Xheneti, Anna Toncheva, Alex Manfrediz: The Expanding Digital Universe: A Forecast of Worldwide Information Growth Through 2010. Framingham 2007.

Gantz et al. / Exploding Universe /

John Gantz, Christopher Chute, Alex Manfrediz, Stephen Minton, David Reinsel, Wolfgang Schlichting, Anna Toncheva: The Diverse and Exploding Digital Universe: An Updated Forecast of Worldwide Information Growth Through 2011. Framingham 2008.

Gantz, Reinsel / Extracting Value /

John Gantz, David Reinsel: Extracting Value from Chaos. Framingham 2011.

Garcia-Molina, Salem / Main Memory Database /

Hector Garcia-Molina, Kenneth Salem: Main Memory Database Systems: An Overview. In: IEEE Transactions on Knowledge and Data Engineering, Band 4, Nr. 6, 1992, S. 509-516.

Gavish, Sheng / Dynamic file migration /

Bezalel Gavish, Olivia Sheng: Dynamic file migration in distributed computer systems. In: Communications of the ACM. Band 33, Nr. 2, 1990, S. 177-189.

Gelb / System-Managed Storage /

Jack Gelb: System-Managed Storage. In: IBM Systems Journal. Band 28, Nr. 1, 1989, S. 77-103.

Giaretta / Digital Preservation /

David Giaretta: Advanced Digital Preservation. Heidelberg 2011.

Gibson, Miller, Long / File-Usage /

Tim Gibson, Ethan Miller, Darrell Long: An Improved Long-Term File-Usage Prediction Algorithm. In: 25th Annual International Conference on Computer Measurement and Performance (CMG 99), Reno 1999, S. 639-648.

Gießelbach / ILM /

Michael Gießelbach: Return-on-Investment durch Information Lifecycle Management. In: 26. DECUS IT-Symposium. Bonn 2003, S. 7-24.

Gilheany / Decline Storage /

Steve Gilheany: The Decline of Magnetic Disc Storage Cost Over the Next 25 Years.

http://www.crescentmeadow.com/document_imaging/pdf/22004p.pdf, Manhattan 2004,

Abruf: 2014-09-15.

Gillet, Mendel / Organic IT /

Frank Gillet, Thomas Mendel: Organic IT: IT-Kosten senken, Unternehmensabläufe beschleunigen. In: Kuhlin, Thielmann (Hrsg.): Real-time Enterprise in der Praxis: Fakten und Praxis. Berlin - Heidelberg - New York 2005; S. 483-502.

Glazer / Measuring Information /

Rashi Glazer: Measuring the value of information: The information-intensive organization. In: IBM System Journal. Band 32, Nr. 1, 1993, S. 99-110.

Goeken / Ansätze der Referenzmodellierung /

Matthias Goeken: Grundlagen und Ansätze einer Referenzmodellierung für Führungsinformationssysteme. Paul Alpar, Ulrich Hasenkamp (Hrsg.): Fachberichte des Instituts für Wirtschaftsinformatik der Philipps-Universität Marburg, Nr. 03/02, Marburg 2003, S. 1-33.

Gohout / Operations Research /

Wolfgang Gohout: Operations Research. Einige ausgewählte Gebiete der linearen und nichtlinearen Optimierung. 4. Auflage, München 2009.

Goldstein / Kontieren /

Elmar Goldstein: Richtig kontieren von A - Z. Das Kontierungslexikon für die Praxis nach DATEV, IKR, BGA. 16. Auflage, Freiburg 2013.

Gonschorrek, Hoffmeister / Planungs- und Entscheidungsprozesse /

Ulrich Gonschorrek, Wolfgang Hoffmeister: Ganzheitliches Management. Planungs- und Entscheidungsprozesse. Berlin 2007.

Gray, Reuter / Transaction Processing /

Jim Gray, Andreas Reuter: Transaction Processing: Concepts and Techniques. San Francisco 1993.

Greasley / Simulation Modelling /

Andrew Greasley: Simulation Modelling for Business. 1. Auflage, Hants-Burlington 2004.

Gregor / Theory of IS /

Shirley Gregor: The Nature of Theory in Information Systems. In: MIS Quarterly, Band 30, Nr. 3, 2006, S. 611-642.

Grobman / ERP-Systeme On Demand /

Jewgenij Grobman: ERP-Systeme On Demand. Chancen, Risiken, Anforderungen, Trends. Hamburg 2008.

Grochla / Organisationstheorie /

Erwin Grochla: Einführung in die Organisationstheorie. Stuttgart 1978.

Guerrero / Modeling and Simulation /

Hector Guerrero: Excel Data Analysis. Modeling and Simulation. 1. Auflage, Heidelberg 2010.

Gugel / Investitionsplanung /

Wolf Gugel: Wissensbasiertes Entscheidungsunterstützungssystem für Investitionsplanung und Investitionskontrolle in der Textilindustrie. Stuttgart 1992.

Gulbins, Seyfried, Zimmermann / Archivierungssysteme /

Jürgen Gulbins, Markus Seyfried, Hans Zimmermann: Elektronische Archivierungssysteme. Berlin u. a. 1993.

Gupta, Sharma, Hsu / Knowledge Based Organizations /

Jatinder Gupta, Sushil Sharma, Jeffrey Hsu: Creating Knowledge Based Organizations. Hershey - London 2004.

Härder, Rahm / Datenbanksysteme /

Theo Härder, Erhard Rahm: Datenbanksysteme: Konzepte und Techniken der Implementierung. 2. Auflage, Berlin 2001.

Hadjiantonis, Pavlou / Self-Management /

Antonis Hadjiantonis, George Pavlou: Policy-Based Self-Management in Wireless Networks. In: Waltenegus Dargie (Hrsg.): Context-Aware Computing and Self-Managing Systems, Boca Raton 2009, S. 201-264.

Hagenloech / Entscheidungslehre /

Thorsten Hagenloech: Grundzüge der Entscheidungslehre. Norderstedt 2009.

Han, Qiu, Huang / Reference Framework /

Shoupeng Han, Xiaogang Qiu, Kedi Huang: DDDAS M&S Reference Framework Based on PADS. In: Koji Koyamada, Shinsuke Tamura, Osamu Ono (Hrsg.): Systems Modeling and Simulation. Proceedings of Asia Simulation Conference 2006, Theory and Applications. Tokyo 2007, S. 83-87.

Hannig / Business Intelligence /

Uwe Hannig: Knowledge Management und Business Intelligence. Berlin - Heidelberg 2002.

Hansen, Neumann / Wirtschaftsinformatik /

Hans Hansen, Gustaf Neumann: Wirtschaftsinformatik 2. Informationstechnik. 9. Auflage, Stuttgart 2005.

Heinen / Führungslehre /

Edmund Heinen: Betriebswirtschaftliche Führungslehre. 2. Auflage, Wiesbaden 1984.

Heinrich / Evaluationsforschung /

Lutz Heinrich: Bedeutung von Evaluation und Evaluationsforschung in der Wirtschaftsinformatik. In: Lutz Heinrich, Irene Häntschel (Hrsg.): Evaluation und Evaluationsforschung in der Wirtschaftsinformatik: Handbuch für Praxis, Lehre und Forschung. München - Wien 2000, S. 7-22.

Heinrich / Systemanalyse /

Gert Heinrich: Allgemeine Systemanalyse. München 2007.

Heinrich, Heinzl, Roithmayr / Wirtschaftsinformatik /

Lutz Heinrich, Armin Heinzl, Friedrich Roithmayr: Wirtschaftsinformatik – Einführung und Grundlegung. 3. Auflage, München 2007.

Heinrich, Stelzer / Informationsmanagement /

Lutz Heinrich, Dirk Stelzer: Informationsmanagement: Grundlagen, Aufgaben, Methoden. 10. Auflage, München 2009.

Heinzel / Nearline Storage /

Marcus Heinzel: Data Life Cycle Management: Mit Nearline Storage Daten parat halten. In: BI-Spektrum. Nr. 4, 2007, S. 31-35.

Heinzl, Schoder, Frank / WI-Journallisten /

Armin Heinzl, Detlef Schoder, Ulrich Frank: WI-Mitteilungen der WKWI und des GI-FB WI. WI-Orientierungslisten. WI-Journalliste 2008 sowie WI-Liste der Konferenzen, Proceedings und Lecture Notes 2008. In: Wirtschaftsinformatik 2, München 2008.

Hennink, Hutter, Bailey / Qualitative Research /

Monique Hennink, Inge Hutter, Ajay Bailey: Qualitative Research Methods. London 2011.

Henze / Stochastik /

Norbert Henze: Stochastik für Einsteiger. Eine Einführung in die faszinierende Welt des Zufalls. 8. Auflage, Wiesbaden 2010.

Herbst / Archivierung in Datenbanksystemen /

Axel Herbst: Anwendungsorientiertes Archivieren in Datenbanksystemen - vertieft am Beispiel von EXPRESS und SDAI. In: Datenbanksysteme in Büro, Technik und Wissenschaft, 1995, S. 194-211.

Hering / Investitionstheorie /

Thomas Hering: Investitionstheorie. 3. Auflage, München 2008.

Hering / Taschenbuch Wirtschaftsingenieure /

Eckbert Hering: Taschenbuch für Wirtschaftsingenieure. 3. Auflage, München 2013.

Herrmann, Spruth / Einführung in z/OS /

Paul Herrmann, Wilhelm Spruth: Einführung in z/OS und OS/390. Web-Services und Internet-Anwendungen für Mainframes. 3. Auflage, München 2012.

Hevner, Chatterjee / Design Research /

Alan Hevner, Samir Chatterjee: Design Research in Information Systems. Theory and Practice. In: Ramesh Sharda, Stefan Voß (Hrsg.): Integrated Series in Information Systems. New York 2010, S. 23-31.

Hevner et al. / Design Science in IS /

Alan Hevner, Salvatore March, Jinsoo Park, Sudha Ram: Design Science in Information Systems Research. In: MIS Quarterly, Band 28, Nr. 1, S. 75-105.

Hiles / Service Level Agreements /

Andrew Hiles: The Complete Guide to IT Service Level Agreements. Aligning IT Service to Business Needs. 3. Auflage, Brookfield 2002.

Hill et al. / Guide Cloud Computing /

Richard Hill: Guide to Cloud Computing. Principles and Practice. 1. Auflage, London 2013.

Hillyer, Silberschatz / Tertiary storage systems /

Bruce Hillyer, Avi Silberschatz: Random I/O scheduling in online tertiary storage systems. In: ACM SIGMOD International Conference on Management of Data, 1996, S. 195-204.

Höfer, Karagiannis / Cloud Computing Services /

Christina Höfer, Georgios Karagiannis: Taxonomy of cloud computing services. In: Journal of Internet Services and Applications, Band 2, Nr. 2, 2011, S. 81-94.

Höllwarth / Cloud Migration /

Tobias Höllwarth: Cloud Migration. 3. Auflage, Heidelberg 2013.

Hölscher / Investition /

Reinhold Hölscher: Investition, Finanzierung und Steuern. München 2010.

Hoffmeister / Investitionsrechnung /

Wolfgang Hoffmeister: Investitionsrechnung und Nutzwertanalyse. Eine Entscheidungsorientierte Darstellung mit vielen Beispielen und Übungen. 2. Auflage, Berlin 2007.

Homburg / Quantitative Betriebswirtschaftslehre /

Christian Homburg: Quantitative Betriebswirtschaftslehre. Entscheidungsunterstützung durch Modelle. Mit Beispielen, Übungsaufgaben und Lösungen. 3. Auflage, Wiesbaden 2000.

Horst / Investition /

Klaus ter Horst: Investition. 2. Auflage, Stuttgart 2009.

House / Evaluating /

Ernest House: Evaluating with Validity. Beverly Hills 1980.

Hübner / Stochastik /

Gerhard Hübner: Stochastik. Eine anwendungsorientierte Einführung für Informatiker, Ingenieure und Mathematiker. 5. Auflage, Wiesbaden 2009.

Hynes, Reynolds, Hauswirth / Context Lifecycle /

Hynes, Reynolds, Hauswirth: A Context Lifecycle for Web-Based Context Management Services. In: Payam Barnaghi, Klaus Moessner, Mirko Presser, Stefan Meissner (Hrsg.): Smart Sensing and Context, Proceedings of the 4th European Conference EuroSSC 2009, Guildford 2009, S. 51-64.

Iannucci, Gupta / Cloud Data Center /

Pietro Iannucci, Manav Gupta: IBM SmartCloud: Building a Cloud Enabled Data Center. 1. Auflage, New York 2013.

Illik / Verteilte Systeme /

Anton Illik: Verteilte Systeme. Architekturen und Software-Technologien. 1. Auflage, Renningen 2007.

Inmon / Data Warehouse /

William Inmon: Building the Data Warehouse. New York 2002.

Inmon / ILM Data Warehousing /

William Inmon: Information Lifecycle Management for Data Warehousing: Matching Technology to Reality. Quebec 2006.

ISACA / Business Framework /

ISACA: COBIT 5. A Business Framework for the Governance and Management of Enterprise IT. Rolling Meadows 2012.

ITGI / Val IT Framework /

IT Governance Institute (ITGI): Enterprise Value: Governance of IT Investments. The Val IT Framework 2.0. Rolling Meadows 2008.

Järvinen / Action Research /

Pertti Järvinen: Action research is similar to design science. In: Quality & Quantity, Band 41, Nr. 1, 2007, S. 37-54.

James / Accelerating ILM /

Derrell James: Accelerating Information Lifecycle Management. A Guide to Planning, Building and Managing a More Cost-effective, Efficient Information Infrastructure.

<http://germany.emc.com/collateral/emc-perspective/h1283-accler-ilm-swp-ldv.pdf>, Hopkinton USA 2004,
Abruf: 2011-04-12.

Jamsa / Cloud Computing /

Kris Jamsa: Cloud Computing: SaaS, PaaS, IaaS, Virtualization, Business Models, Mobile, Security and More. Burlington 2012.

Janko / Informationswirtschaft /

Wolfgang Janko: Informationswirtschaft 1. Grundlagen der Informatik für die Informationswirtschaft. 2. Auflage, Springer-Verlag, Berlin 1998.

Janukowicz / Solid State Storage /

Jeff Janukowicz: How Enterprise Architectures Must Adapt to Solid State Storage. Farmingham 2011.

Jensen, Reed / Activity Computer Environment /

David Jensen, Daniel Reed: File Archive Activity in a Supercomputer Environment. In: Proceedings of the 7th international conference on Supercomputing, Tokyo 1993, S. 387-396.

Jeong et al. / Effective Storage /

Seung Jeong, Yoosin Kim, Imran Ghani, Jae Kim: A New Database Archiving Approach for Effective Storage and Data Management: A Case Study of Data Warehouse Project in a Korean Bank. In: International Journal of Advances in Soft Computing and its Applications, Band 6, Nr. 3, 2014, S. 30-46.

Jewell et al. / Performance and Capacity /

Dave Jewell, Ricardo Barros, Stefan Diederichs, Lydia Duijvestijn, Michael Hammersley, Arindam Hazra, Corneliu Holban, Yan Li, Osai Osaigbovo, Andreas Plach, Ivan Portilla, Mukerji Saptarshi, Harinder Seera, Elisabeth Stahl, Clea Zolotow: Performance and Capacity Implications for Big Data. 1. Auflage, New York 2014.

Jin et al. / Cloud Types and Services /

Hai Jin, Shadi Ibrahim, Tim Bell, Wei Gao, Dachuan Huang, Song Wu: Cloud Types and Services. In: Borko Furht, Armando Escalante (Hrsg.): Handbook of Cloud Computing, New York 2010, S. 335-356.

Jo et al. / Virtual Disk /

Heeseung Jo, Youngjin Kwon, Hwanju Kim, Euiseong Seo, Joonwon Lee, Seungryoul Maeng: SSD-HDD-hybrid virtual disk in consolidated environments. In: Hai-Xiang Lin, Sips Henk, Epema Dick (Hrsg.): Euro-Par'09 Proceedings of the 2009 international conference on Parallel processing, Berlin et al. 2009, S. 375-384.

Johnson, Agrawal / Law and Technology /

Christopher Johnson, Rakesh Agrawal: Intersections of Law and Technology in Balancing Privacy Rights with Free Information Flow. In: Hafner, O'Rourke (Hrsg.): 4th International Conference on Law and Technology, Cambridge 2006.

Jonscher / Economic Study IT /

Charles Jonscher: An Economic Study of the Information Technology Revolution. In: Thomas Allen, Michael Morton (Hrsg.): Information Technology and the Corporation of the 1990s. Research Studies. Oxford 1994, S. 5-42.

Jung / Controlling /

Hans Jung: Controlling. 2. Auflage, München 2007.

Kaariainen, Valimaki / Application Lifecycle Management /

Jukka Kaariainen, Antti Valimaki: Impact of Application Lifecycle Management - A Case Study. In: Kai Mertins, Rainer Ruggaber, Keith Popplewell, Xiaofei Xu (Hrsg.): Enterprise Interoperability III. New Challenges and Industrial Approaches, London 2008, S. 54-69.

Kaarst-Brown, Kelly / IT-Governance /

Michelle Kaarst-Brown, Shirley Kelly: IT-Governance and Sarbanes Oxley: The latest sales pitch or real challenges for the IT Function? In: Proceedings of the 38th Hawaii International Conference on System Sciences, Big Islands 2005.

Kaiser Smolnik / ILM für Dokumente /

Michael Kaiser, Stefan Smolnik: Information Lifecycle Management für Dokumente. In: Stefan Walter, Gernot Kaiser (Hrsg.): Dokumentenlogistik. Theorie und Praxis. Berlin - Heidelberg 2014, S. 89-108.

Kaiser, Smolnik, Riempp / ILM-Framework /

Michael Kaiser, Stefan Smolnik, Gerold Riempp: Konzeption eines Information-Lifecycle-Management-Frameworks im Dokumenten-Management-Kontext. In: Martin Bichler, Thomas Hess, Helmut Krcmar, Ulrike Lechner, Florian Matthes, Arnold Picot, Benjamin Speitkamp, Petra Wolf (Hrsg.): Multikonferenz Wirtschaftsinformatik 2008. Berlin 2008, S. 483-494.

Kaiser, Smolnik, Riempp / Verbesserte Compliance ILM /

Michael Kaiser, Stefan Smolnik, Gerold Riempp: Verbesserte Compliance durch Information Lifecycle Management. In: HMD Praxis der Wirtschaftsinformatik, Band 45, Nr. 5, 2008, S. 30-38.

Kanakamedala, Kaplan, Srinivasaraghavan / Data storage /

Kishore Kanakamedala, James Kaplan, Rajesh Srinivasaraghavan: A smarter approach to data storage. In: The McKinsey Quarterly, 2007, S. 1-4.

Kaplan, Rezek, Sprague / Information Protection Cloud /

James Kaplan, Chris Rezek, Kara Sprague: Protecting information in the cloud. In: McKinsey on Business Technology, Nr. 28, San Francisco 2012, S. 12-19.

Kaplan, Roy, Srinivasaraghavan / Demand data storage /

James Kaplan, Rishi Roy, Rajesh Srinivasaraghavan: Meeting the demand for data storage. In: The McKinsey Quarterly, 2008, S. 1-10.

Keeney, Raiffa / Multiple Objectives /

Ralph Keeney, Howard Raiffa: Decisions with Multiple Objectives: Preferences and Value Tradeoffs. Cambridge University-Press, 1993.

Kemmettmüller, Bogensberger / Kostenrechnung /

Wolfgang Kemmettmüller, Stefan Bogensberger: Handbuch der Kostenrechnung. Das Grundlagenwerk zu Kostenrechnung und Kostenmanagement. 8. Auflage, Wien 2004.

Kersten, Klett / IT-Security /

Heinrich Kersten, Gerhard Klett: Der IT Security Manager. Wiesbaden 2005.

Kersten, Reuter, Schröder / IT-Sicherheitsmanagement /

Heinrich Kersten, Jürgen Reuter, Klaus-Werner Schröder: IT-Sicherheitsmanagement nach ISO 27001 und Grundschutz. Der Weg zur Zertifizierung. 4. Auflage, New York 2013.

Killich, Maicher, Pfeiffer / IT-Value /

Jürgen Killich, Michael Maicher, Tony Pfeiffer: IT-Value- und Portfoliomanagement - Definitionen und Konzepte. In: Sebastian Kammerer, Michael Lang, Michael Amberg (Hrsg.): IT-Projektmanagement Methoden. Best Practices von Scrum bis PRINCE2. Düsseldorf 2012, S. 53-79.

Kofler, Nebelo / Excel Programmieren /

Michael Kofler, Ralf Nebelo: Excel Programmieren. Abläufe automatisieren, Apps und Anwendungen entwickeln. 1. Auflage, Kösel 2013.

Kolonko / Stochastische Simulation /

Michael Kolonko: Stochastische Simulation. Grundlagen, Algorithmen und Anwendungen. 1. Auflage, Wiesbaden 2008.

Königer, Reithmayer / Management Informationen /

Paul König, Walter Reithmayer: Management unstrukturierter Informationen: Wie Unternehmen die Informationsflut beherrschen können. Frankfurt a. M. 1998.

Komorowski / Cost per Gigabyte /

Matthew Komorowski: Hard Drive Costs per Gigabyte.
<http://ns1758.ca/winch/winchest.html>, o. O. 2012,
Abruf: 2013-01-05.

Kosler, Matthesius, Stelzer / Klassifizierung von Informationen /

Monique Kosler, Mauricio Matthesius, Dirk Stelzer: Ein Konzept zur automatisierten Klassifizierung von Informationen für das Information Lifecycle Management - dargestellt am Beispiel des SAP NetWeaver Business Intelligence. In: Barbara Dinter, Robert Winter, Peter Chamoni, Norbert Gronau, Klaus Turovski (Hrsg.): DW2008. Synergien durch Integration und Informationslogistik. Bonn 2008, S. 129-145.

Krcmar / Informationsmanagement /

Helmut Krcmar: Informationsmanagement. 4. Auflage, Berlin - Heidelberg - New York 2006.

Krieger / Kostenrechnung /

Ralf Krieger: Betriebsindividuelle Gestaltung der Kostenrechnung. Ein Beitrag zur Weiterentwicklung der situativen Kostenrechnungstheorie unter besonderer Berücksichtigung der Industriebetriebe. Berlin 1995.

Kriesi / Rechtliche Aspekte von SaaS /

Lilian Snaidero Kriesi: Die rechtlichen Aspekte von SaaS. In: Swiss IT Magazine, Nr.3, 2012, S. 40-41.

Krishnan / Big Data /

Krish Krishnan: Data Warehousing in the Age of Big Data. Waltham 2013.

Krüger / Marktfähige Produkte /

Carsten Krüger: Die VEAG und ihr IT-Partner – Von pauschalierten Dienstleistungen zu marktfähigen Produkten. In: Martin Bernhard, Winfried Lewandowski, Hartmut Mann (Hrsg.): Service-Level-Management in der IT. Wie man erfolgskritische Leistungen definiert und steuert. 5. Auflage, Düsseldorf 2004, S. 135-152.

Küll, Stähly / Simulationsexperimente /

Roland Küll, Paul Stähly: Zur Planung und effizienten Abwicklung von Simulationsexperimenten. In: Jörg Biethahn, Wilhelm Hummeltenberg, Bernd Schmidt, Paul Stähly, Thomas Witte (Hrsg.): Simulation als betriebliche Entscheidungshilfe. State of the Art und neuere Entwicklungen, Heidelberg 1999, S. 1-21.

Küspert, Schaarschmidt, Herbst / Archiv Transaktionen /

Klaus Küspert, Ralf Schaarschmidt, Axel Herbst: Archiv Transaktionen: Ein Ansatz für asynchrones, transaktionsgesichertes Archivieren großer Datenmengen in Datenbanksystemen. In: ITG-Fachbericht zur 4. ITG/GI/GMA-Fachtagung Softwaretechnik in Automation und Kommunikation - Rechnergestützte Teamarbeit (STAK). München 1996, S. 195-211.

Kuhn, Hellingrath / SCM /

Axel Kuhn, Bernd Hellingrath: Supply Chain Management: Optimierte Zusammenarbeit in der Wertschöpfungskette. Berlin - Heidelberg - New York 2002.

Kummer, Grün, Jammernegg / Produktion und Logistik /

Sebastian Kummer, Oskar Grün, Werner Jammernegg: Grundzüge der Beschaffung, Produktion und Logistik. 2. Auflage, München 2009.

Kußmaul / Betriebswirtschaftslehre /

Heinz Kußmaul: Betriebswirtschaftslehre für Existenzgründer. Grundlagen mit Fallbeispielen und Fragen der Existenzgründungspraxis. 6. Auflage, München 2008.

Laahs, McKenna, Vanamo / Sharepoint-Technologien /

Kevin Laahs, Emer McKenna, Veli Vanamo: Microsoft SharePoint-Technologien. Planung, Design und Implementierung von Windows SharePoint Services und SharePoint Portal Server. 1. Auflage, München 2005.

Lampe / Green-IT /

Frank Lampe: Green-IT, Virtualisierung und Thin-Clients. Mit neuen IT-Technologien Energieeffizienz erreichen, die Umwelt schonen und Kosten sparen. Wiesbaden 2010.

Lassmann / Wirtschaftsinformatik /

Wolfgang Lassmann: Wirtschaftsinformatik: Nachschlagewerk für Studium und Praxis. Wiesbaden 2006.

Laudon, Laudon, Schoder / Wirtschaftsinformatik /

Kenneth Laudon, Jane Laudon, Detlef Schoder: Wirtschaftsinformatik. Eine Einführung. 2. Auflage, München 2010.

Lawrie, Randal; Barton / File Migration /

Duncan Lawrie, Michael Randal, Richard Barton: Experiments with Automatic File Migration. In: IEEE Computer, Nr. 7, 1982, S. 45-55.

Lawton / Moving OS to the Web /

George Lawton: Moving the OS to the Web. In: Computer, Band 41, Nr. 3, 2008, S. 16-19.

Lee et al. / ILM with RFID /

Ju Hyun Lee, Jeonghwa Song, Kunsoo Oh, Ning Gu: Information Lifecycle Management with RFID for material control on construction sites. In: Journal of Advanced Engineering Informatics, Band 27, Nr. 1, 2013, S. 108-119.

Lehner / Wissensmanagement /

Franz Lehner: Wissensmanagement. Grundlagen, Methoden und technische Unterstützung. 5. Auflage, München 2014.

Lehner, Sattler / Web-Scale Cloud /

Wolfgang Lehner, Kai-Uwe Sattler: Web-Scale Data Management for the Cloud. New York 2013.

Lenhardt et al. / Data Lifecycle /

Christopher Lenhardt, Stanley Ahalt, Brian Blanton, Laura Christopherson, Ray Idaszak: Data Management Lifecycle and Software Lifecycle Management in the Context of Conducting Science. In: Journal of Open Research Software, Band 2, Nr. 1, 2014, S. 1-4.

Lepsky, Siepmann, Zimmermann / Automatische Indexierung /

Klaus Lepsky, Jörg Siepmann, Andrea Zimmermann: Automatische Indexierung für Online-Kataloge: Ergebnisse eines Retrievaltests. In: Zeitschrift für Bibliothekswesen und Bibliographie, Band 43, Nr. 1, 1996, S. 47-56.

Levy / NSA killed the Internet /

Steven Levy: How the NSA Almost killed the Internet.
<http://www.wired.com/threatlevel/2014/01/how-the-us-almost-killed-the-internet/>, o.O 2014,
Abruf: 2014-01-08.

Li et al. / Authentication Cloud /

Hongwei Li, Yuanshun Dai, Ling Tian, Haomiao Yang: Identity-Based Authentication for Cloud Computing. In: Martin Jaatun, Gansen Zhao, Chunming Rong (Hrsg.): Cloud Computing: Proceedings of the First International Conference, CloudCom 2009, Beijing 2009, S. 157-166.

Liedloff, Bromberger / Datenhaltung /

Thilo Liedloff, Heiko Bromberger: Informationen für morgen aus Systemen von gestern? - Der IBM Mainframe im Mittelpunkt zentraler Datenhaltung im Jahr 2020. In: Frank Keuper, Fritz Neumann (Hrsg.): Wissens- und Informationsmanagement. 1. Auflage, Wiesbaden 2009, S. 571-594.

Lin, Devine / Networks in Cloud Computing /

Geng Lin, Mac Devine: The Role of Networks in Cloud Computing. In: Borko Furht, Armando Escalante (Hrsg.): Handbook of Cloud Computing. New York 2010, S. 65-82.

Linthicum / Cloud Contracts /

David Linthicum: Today's cloud contracts are driving away enterprise adoption. As cloud providers push 'take it or leave it' contracts, many businesses push back – or avoid the cloud altogether.
<http://www.infoworld.com/d/cloud-computing/todays-cloud-contracts-are-driving-away-enterprise-adoption-211624>, o. O. 2013,
Abruf: 2014-01-08.

Lippold / Unternehmensberatung /

Dirk Lippold: Die Unternehmensberatung. Von der strategischen Konzeption zur praktischen Umsetzung. Wiesbaden 2013.

Lippold, Stelzer, Konrad / Sicherheitskonzepte /

Heiko Lippold, Dirk Stelzer, Peter Konrad: Sicherheitskonzepte und ihre Verknüpfung mit Sicherheitsstrategie und Sicherheitsmanagement. In: Wirtschaftsinformatik, Nr. 4, 1992, S. 367-377.

Liu, Ngoc, Laurent / Cost Framework /

Tianxiao Liu, Tuyet Ngoc, Dominique Laurent: Cost Framework for a Heterogeneous Distributed Semi-structured Environment. In: Kevin Chang, Wei Wang, Lei Chen, Clarence Ellis, Ching Hsu, Ah Tsoi, Haixun Wang (Hrsg.): Proceedings of the International Workshop on Database Management and Application over Networks. Berlin, Heidelberg 2007, S. 1-11.

Liu, Wang, Quan / ILM Architecture /

Huiyan Liu, Xueying Wang, Qqizhe Quan: Research on the Enterprise Model of Information Lifecycle Management Based on Enterprise Architecture. In: Proceedings of the 2009 9th International Conference on Hybrid Intelligent Systems, 3. Band. Shenyang 2009, S. 165-169.

Liu, Li, Zhang / Dispatching Methods /

Baoliang Liu, Jianzhong Li, Yanqiu Zhang: Optimal Data Dispatching Methods in Near-Line Tertiary Storage System. In: Qing Li, Guoren Wang, Ling Feng (Hrsg.): Proceedings of the 5th International Conference: Advances in Web-Age Information Management. Dalian, China 2004, S. 690-695.

Loos et al. / In-Memory /

Peter Loos, Jens Lechtenbörger, Gottfried Vossen, Alexander Zeier, Jens Krüger, Jürgen Müller, Wolfgang Lehner, Donald Kossmann, Benjamin Fabian, Oliver Günther, Robert Winter: In-Memory-Datenmanagement in betrieblichen Anwendungssystemen. In: Wirtschaftsinformatik, Jahrgang 53, Nr. 6, 2011, S. 383-390.

Loshin / Enterprise Knowledge Management /

David Loshin: Enterprise Knowledge Management. San Diego - London – San Francisco 2001.

Lotlikar, Mohania / Policies ILM /

Rohit Lotlikar, Mukesh Mohania: Adaptive policies in Information Lifecycle Management. In: Stephane Bressan, Josef Küng, Roland Wagner (Hrsg.): Proceedings of the 17th international conference on Database and Expert Systems Applications, Krakau 2006, S. 612-621.

Lundberg / Lifecycle Management /

Jeff Lundberg: Understanding Data Lifecycle Management.
<http://sysdoc.doors.ch/VERITAS/4434.pdf>, Mountain View 2003,
Abruf: 2011-04-13.

MacKenzie et al. / Reference Model SOA /

Matthew MacKenzie, Ken Laskey, Francis McCabe, Peter Brown, Rebekah Metz: Reference Model for Service Oriented Architecture 1.0.
<http://docs.oasis-open.org/soa-rm/v1.0/soa-rm.pdf>, o. O. 2006,
Abruf: 2013-06-06.

Magoulès, Pan, Teng / Data Intensive Computing /

Frédéric Magoulès, Jie Pan, Fei Teng: Cloud Computing. Data-Intensive Computing and Scheduling. Boca Raton 2013.

Mahmood / Cloud Computing /

Zaigham Mahmood: Cloud Computing. Methods and Practical Approaches. 1. Auflage, London et al. 2013.

Mahmood, Saeed / Frameworks for the Cloud /

Zaigham Mahmood, Saqib Saeed: Software Engineering Frameworks for the Cloud Computing Paradigm. 1. Auflage, London 2013.

Maier, Hädrich, Peinl / Knowledge Infrastructures /

Ronald Maier, Thomas Hädrich, René Peinl: Enterprise Knowledge Infrastructures. 2. Auflage, Berlin - Heidelberg - New York 2009.

Mall / Real-Time Systems /

Rajib Mall: Real-Time Systems. Theory and Practice. 2. Auflage, New Delhi 2008.

Mansmann / Breitband /

Urs Mansmann: Breitband rasant. Internetanschlüsse werden immer schneller. In: c't Magazin Nr. 25, 2010, S. 92.

Marx Gómez, Junker, Odebrecht / IT-Controlling /

Jorge Marx Gómez, Horst Junker, Stefan Odebrecht / IT-Controlling. Strategien, Werkzeuge, Praxis. Göttingen 2009.

Masak / IT-Alignment /

Dieter Masak: IT-Alignment. Berlin - Heidelberg - New York 2006.

Mather, Kumaraswamy, Latif / Cloud Security /

Tim Mather, Subra Kumaraswamy, Shahed Latif: Cloud Security and Privacy. An Enterprise Perspective on Risks and Compliance. 1. Auflage, Sebastopol 2009.

Matschke, Hering, Klingelhöfer / Finanzplanung /

Manfred Matschke, Thomas Hering, Heinz Klingelhöfer: Finanzanalyse und Finanzplanung. München - Wien 2002.

Matthesius / ILM-Konzept Energiekonzern /

Mauricio Matthesius: Entwicklung eines Information Lifecycle Management-Konzepts in einer komplexen Data Warehouse-Landschaft am Beispiel eines Energiekonzerns. Diplomarbeit vorgelegt an der Fakultät für Wirtschaftswissenschaften, Institut für Wirtschaftsinformatik der Technischen Universität Ilmenau, Ilmenau 2007.

Matthesius, Stelzer / Analyse und Vergleich Konzepte im ILM /

Mauricio Matthesius, Dirk Stelzer: Analyse und Vergleich von Konzepten zur automatisierten Informationsbewertung im Information Lifecycle Management. In: Martin Bichler, Thomas Hess, Helmut Krcmar, Ulrike Lechner, Florian Matthes, Arnold Picot, Benjamin Speitkamp, Petra Wolf (Hrsg.): Multikonferenz Wirtschaftsinformatik 2008. Berlin 2008, S. 471-482.

McGrath / Understanding PaaS /

Michael McGrath: Understanding PaaS. Unleash the Power of Cloud Computing. Sebastopol 2012.

McKinsey / Challenge Big Data /

McKinsey: The challenge and opportunity of big data. In: The McKinsey Quarterly, 2011, S. 1-3.

McLeod, Jordan / Systems Development /

Raymond McLeod, Eleanor Jordan: Systems Development. A Project Management Approach. Austin 2001.

McMillan / Role Cloud Computing in BigData /

Liz McMillan: The Role of Cloud Computing in BigData.
<http://www.sys-con.com/node/2585004>, o. O. 2013,
Abruf: 2013-03-21.

McPeak / Simulation Noise /

John McPeak: Simulation Noise and the Estimation of Land Use Decision in Kenya. In: Riccardo Scarpa, Anna Alberini (Hrsg.): Applications of Simulation Methods in Environmental and Resource Economics. Dordrecht 2005, S. 355-372.

Mehlan / Controlling /

Axel Mehlan: Praxishilfen Controlling. Die besten Controlling-Instrumente mit Excel. München 2007.

Meinel et al. / Cloud-Speicherdienste /

Christoph Meinel, Maxim Schnjakin, Tobias Metzke, Markus Freitag: Anbieter von Cloud Speicherdiensten im Überblick. Technischer Bericht Nr. 84, Hasso Plattner Institut für Softwaresystemtechnik der Universität Potsdam, Potsdam 2014.

Mell, Grance / Definition Cloud Computing /

Peter Mell, Timothy Grance: The NIST Definition of Cloud Computing. Recommendations of the National Institute of Standards and Technology. <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>, Gaithersburg 2011, Abruf: 2013-02-07.

Melzer / Architekturen mit Web Services /

Ingo Melzer: Service-orientierte Architekturen mit Web Services. 4. Auflage, Heidelberg 2010.

Mensch / Investition /

Gerhard Mensch: Investition: Investitionsrechnung in der Planung und Beurteilung von Investitionen. München-Wien-Oldenbourg 2002.

Mertens / Evaluation /

Donna Mertens: Institutionalizing Evaluation in the United States of America. In: Reinhard Stockmann (Hrsg.): Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder. 3. Auflage, Münster 2006, S. 47-63.

Mertens et al. / Grundzüge Wirtschaftsinformatik /

Peter Mertens, Freimut Bodendorf, Wolfgang König, Arnold Picot, Matthias Schumann, Thomas Hess: Grundzüge der Wirtschaftsinformatik. 4. Auflage, Berlin - Heidelberg - New York 2004.

Merrill, Thanhardt / Mass Storage /

John Merrill, Eric Thanhardt: Early Experience with Mass Storage on a UNIXBased Supercomputer. In: Proceedings of the 10th IEEE Symposium on Mass Storage Systems. Monterey 1990, S. 117-121.

Mesnier et al. / File classification /

Michael Mesnier, Eno Thereska, Daniel Ellard, Gregory Ganger, Marco Seltzer: File classification in self-* storage systems. In: Proceedings of the 1st International Conference on Autonomic Computing (ICAC-04). New York 2004, S. 44-51.

Meyer / Betriebswirtschaftliche Kennzahlen /

Claus Meyer: Betriebswirtschaftliche Kennzahlen und Kennzahlen-Systeme. 6. Auflage, Paderborn 2011.

Michelsen, English / Service Virtualization /

John Michelsen, Jason English: Service Virtualization. Reality is Overrated. How to Win the Innovation Race by Virtualizing Everything. New York 2012.

Middleton / Data-Intensive Technologies for Cloud Computing /

Anthony Middleton: Data-Intensive Technologies for Cloud Computing. In: Borko Furht, Armando Escalante (Hrsg.): Handbook of Cloud Computing. New York et al. 2010, S. 83-136.

Miebach / Handlungstheorie /

Bernhard Miebach: Soziologische Handlungstheorie. Eine Einführung. Wiesbaden 2010.

Miller, Katz / Analysis File Migration /

Ethan Miller, Randy Katz: An Analysis of File Migration in a UNIX Supercomputing Environment. In: USENIX Winter 1993 Conference. San Diego 1993, S. 421-434.

Miller, Katz / Behavior Supercomputer /

Ethan Miller, Randy Katz: Analyzing the I/O Behavior of Supercomputer Applications. In: Proceedings of the 11th IEEE Symposium on Mass Storage Systems. Monterey 1991, S. 51-55.

Minoli / Multicast Applications /

Daniel Minoli: Linear and Non-Linear Video and TV Applications. Using IPv6 and IPv6 Multicast. 1. Auflage, Hoboken 2012.

Mirchandani / Technology Innovations /

Vinnie Mirchandani: The New Polymath. Profiles in Compound Technology-Innovations. 1. Auflage, Hoboken 2010.

Möller / Lifecycle models /

Knud Möller: Lifecycle models of data-centric systems and domains: The abstract data lifecycle model. In: Semantic Web Journal, Band 4, Nr. 1, S. 67-88.

Möhrle, Isenmann / Zukunftsstrategien /

Martin Möhrle, Ralf Isenmann: Technologie-Roadmapping: Zukunftsstrategien für Technologieunternehmen. 3. Auflage, Berlin - Heidelberg - New York 2007.

Moens / Automatic Indexing /

Marie-Francine Moens: Automatic Indexing and Abstracting of Document Texts. In: Artificial, Intelligence and Law. Band 3, Nr. 4, 2000, S. 343-347.

Mohanty, Jagadeesh, Srivatsa / Big Data Imperatives /

Soumendra Mohanty, Madhu Jagadeesh, Harsha Srivatsa: Big Data Imperatives. Enterprise Big Data Warehouse, BI Implementations and Analytics. Berkley 2013.

Mont / Identity Lifecycle Management /

Marco Mont: Privacy-Aware Identity Lifecycle Management. In: Digital Privacy, Lecture Notes in Computer Science, Band 6545, 2011, S. 397-426.

Mont / Privacy-aware ILM /

Marco Mont: On Privacy-aware Information Lifecycle Management in Enterprises: Setting the Context. In: Paulus, Pohlmann, Reimer (Hrsg.): ISSE 2006 - Securing Electronic Business Processes. Vieweg-Verlag, Wiesbaden 2006, S. 405-415.

Mont, Beato / Obligation Policies /

Marco Mont, Filipe Beato: On Parametric Obligation Policies: Enabling Privacy-aware Information Lifecycle Management in Enterprises. In (IEEE Computer Society): Proceedings of the 8th IEEE International Workshop on Policies for Distributed Systems and Networks, Washington 2007, S. 51-55.

Mont et al. / Privacy Policy Enforcement /

Marco Mont, Robert Thyne, Pete Bramhall, Kwok Chan: Privacy Policy Enforcement in Enterprises. Addressing Regulatory Compliance and Governance Needs. In: Sachar Paulus, Norbert Pohlmann, Helmut Reimer (Hrsg.): ISSE 2005 Securing Electronic Business Processes. Highlights of the Information Security Solutions Europe 2005 Conference, Wiesbaden 2005, S. 137-147.

Moody, Walsh / Measuring Information /

Moody, D.; Walsh, P.: Measuring the value of information: An asset valuation approach. In: Proceedings of the 7th European Conference on Information Systems. Kopenhagen - Frederiksberg 1999, S. 496-512.

Moore / ILM /

Fred Moore: Information Lifecycle Management. Horison Information Strategies, Melbourne 2004.

Müller / IT-Sicherheit /

Klaus Müller: IT-Sicherheit mit System. 3. Auflage, Wiesbaden 2007.

Münch / Datenschutz /

Peter Münch: Technisch-organisatorischer Datenschutz. Leitfaden für Praktiker. 4. Auflage, Heidelberg et al. 2010.

Mugg / Storage für KMU /

Christian Mugg: Storage as a Service für KMU. In: Swiss IT Magazine, Nr.3, 2012, S. 32-34.

Nebl / Produktionswirtschaft /

Theodor Nebl: Produktionswirtschaft: 7. Auflage, München 2011.

Nehfort / IT-Lösungen /

Andreas Nehfort: Qualitätsmanagement für IT-Lösungen. In: Ernst Tiemeyer (Hrsg.): Handbuch IT-Management. Konzepte, Methoden, Lösungen und Arbeitshilfen für die Praxis. 5. Auflage, München 2013, S. 493-549.

Nissen, Biethahn / Stochastische Optimierung /

Volker Nissen, Jörg Biethahn: Ein Beispiel zur stochastischen Optimierung mittels Simulation und einem genetischen Algorithmus. In: Jörg Biethahn, Wilhelm Hummeltenberg, Bernd Schmidt, Paul Stähly, Thomas Witte (Hrsg.): Simulation als betriebliche Entscheidungshilfe. State of the Art und neuere Entwicklungen, Heidelberg 1999, S. 108-125.

Nissen, Matthesius / Forecast Model Selection /

Volker Nissen, Mauricio Matthesius: Automatic Forecast Model Selection in SAP Business Information Warehouse BPS - an Empirical Investigation. In: Udo Bankhofer, Peter Gmilkowsky, Volker Nissen, Dirk Stelzer (Hrsg.): Ilmenauer Beiträge zur Wirtschaftsinformatik. Technischer Report Nr. 2006-04, Technische Universität Ilmenau 2006.

Nissen, Propach / Optimization Noise /

Volker Nissen, Jörn Propach: On the Robustness of Population-Based Versus Point-Based Optimization in the Presence of Noise. In: IEEE Transactions on Evolutionary Computation, Band 2, Nr. 3, 1998, S. 107-119.

NLNZ / Metadata Extraction Tool /

National Library of New Zealand: Metadata Extraction Tool.
[http:// http://meta-extractor.sourceforge.net](http://meta-extractor.sourceforge.net), New Zealand 2007,
Abruf: 2014-01-26.

Nohr / Wissensmanagement /

Holger Nohr: Wissensmanagement. Wie Unternehmen ihre wichtigste Ressource erschließen und teilen. Göttingen 2000.

Null, Lobur / Computer Organization /

Linda Null, Julia Lobur: The Essentials of Computer Organization and Architecture. 4. Auflage, Burlington 2014.

Nunamaker, Chen, Purdin / Systems Development /

Jay Nunamaker, Minder Chen, Titus Purdin: Systems Development in Information Systems Research. In: Journal of Management Information Systems, Band 7, Nr. 3, 1991, S. 89-106.

Oates et al. / Spring und Hibernate /

Richard Oates: Spring & Hibernate. Eine praxisbezogene Einführung. 2. Auflage, München-Wien 2008.

Obermeier, Gasper / Investitionsrechnung /

Thomas Obermeier, Richard Gasper: Investitionsrechnung und Unternehmensbewertung. München 2008.

OECD / Growth and Innovation /

Organisation for Economic Co-Operation and Development: Supporting Investment in Knowledge Capital, Growth and Innovation. Paris 2013.

Onwubiko / Security Issues Cloud /

Cyril Onwubiko: Security Issues to Cloud Computing. In: Nick Antonopoulos, Lee Gillam (Hrsg.): Cloud Computing. Principles, Systems and Applications, London et al. 2010, S. 271-288.

Oppenheim, Stenson, Wilson / Information as an asset /

Charles Oppenheim, Joan Stenson, Richard Wilson: The attributes of information as an asset. In: New Library World, Band 102, Nr. 11/12, 2001, S. 458-464.

Os, Bressan / Cloud Real-Time Application /

Marcelo Os, Graca Bressan: A Community Cloud for a Real-Time Financial Application - Requirements, Architecture and Mechanisms. In: Xian-he Sun, Wenyu Qu, Ivan Stojmenovic, Wanlei Zhou, Zhiyang Li, Hua Guo, Geyong Min, Tingting Yang, Yulei Wu, Lei Liu (Hrsg.): Algorithms and Architectures for Parallel Processing, 14th International Conference, Dalian 2014, S. 364-377.

Oßmann / Daten ILM /

Kirsten Oßmann: Automatisierte Bewertung von Daten im SAP Business Information Warehouse im Rahmen des Information Lifecycle Managements. Diplomarbeit vorgelegt an der Fakultät für Wirtschaftswissenschaften, Institut für Wirtschaftsinformatik der Technischen Universität Ilmenau, Ilmenau 2008.

Ott / Wirtschaftlichkeitsanalyse /

Hans Jürgen Ott: Wirtschaftlichkeitsanalyse von EDV-Investitionen mit dem WARS-Modell am Beispiel der Einführung von CASE. In: Wirtschaftsinformatik, Band 35, Nr. 6, 1993, S. 522-531.

Pallickara, Pallickara, Pierce / Data Management in the Cloud /

Sangmi Pallickara, Shrideep Pallickara, Marlon Pierce: Scientific Data Management in the Cloud: A Survey of Technologies, Approaches and Challenges. In: Borko Furht, Armando Escalante (Hrsg.): Handbook of Cloud Computing, New York et al. 2010, S. 517-534.

Pamieri, Pardi / Cloud Computing /

Francesco Pamieri, Silvio Pardi: Enhanced Network Support for Scalable Computing Clouds. In: Nick Antonopoulos, Lee Gillam (Hrsg.): Cloud Computing. Principles, Systems and Applications, London et al. 2010, S. 127-144.

Park / Privacy Policy Grid /

Namje Park: Implementation of Privacy Policy-based Protection System in BEMS based Smart Grid Service. In: International Journal of Smart Home, Band 7, Nr. 6, 2013, S. 91-100.

Park, Kim / Policy energy management /

Namje Park, Marie Kim: Implementation of load management application system using smart grid privacy policy in energy management service environment. In: Cluster Computing, Band 17, Nr. 3, 2014, S. 653-664.

Patrascu et al. / Autonomic computing /

Relu Patrascu, Craig Boutilier, Rajarshi Das, Jeffrey Kephart, Gerald Tesauero, William Walsh: New approaches to optimization and utility elicitation in autonomic computing. In: Proceedings of the National Conference on Artificial Intelligence. Pittsburgh 2005, S. 140-145.

Peffers et al. / Research Evaluation /

Ken Peffers, Marcus Rothenberger, Tuure Tuunanen, Reza Vaezi: Design Science Research Evaluation. In: Ken Peffers, Marcus Rothenberger, Bill Kuechler (Hrsg.): Design Science Research in Information Systems. Advances in Theory and Practice. DESRIST'12 Proceedings, of the 7th international conference on Design Science Research in Information Systems, Band 7286, 2012, S. 398-410.

Peffers et al. / Research Methodology /

Ken Peffers, Marcus Rothenberger, Tuure Tuunanen, Samir Chatterjee: A Design Science Research Methodology for Information Systems Research. In: Journal of Management Information Systems, Band 24, Nr. 3, 2008, S. 45-77.

Peters / Organisationsgestaltung /

Heiko Peters: Organisationsgestaltung durch explizite Verhaltensnormen. Eine agentenbasierte Simulation zur Organisation dezentraler Informationsverarbeitung. Schriften zu Management, Organisation und Information. Band 35, München 2012.

Peters, Brühl, Stelling / Betriebswirtschaftslehre /

Sönke Peters, Rolf Brühl, Johannes Stelling: Betriebswirtschaftslehre. 12. Auflage, München 2005.

Petrocelli / Information Lifecycle Management /

Tom Petrocelli: Data Protection and Information Lifecycle Management. New York u.a. 2005.

Pfnür, Schetter, Schöbener / Risikomanagement /

Andreas Pfnür, Christoph Schetter, Henning Schöbener: Risikomanagement bei Public Private Partnerships. Heidelberg et al. 2010.

Picot, Reichwald, Wigand / Unternehmung /

Arnold Picot, Ralf Reichwald, Rolf Wigand: Die grenzenlose Unternehmung. 5. Auflage, Wiesbaden 2003.

Piro, Gebauer / Daten- und Informationsqualität /

Andrea Piro, Marcus Gebauer: Definition von Datenarten zur konsistenten Kommunikation in Unternehmen. In: Knut Hildebrand, Marcus Gebauer, Holger Hinrichs, Michael Mielke (Hrsg.): Daten- und Informationsqualität. Auf dem Weg zur Information Excellence. Wiesbaden 2008, S. 143-156.

Plattner / Database Approach /

Hasso Plattner: A common database approach to OLTP and OLAP using an in-memory column database. In: Proceedings of the ACM SIGMOD International Conference on Management of Data, SIGMOD 2009, Rhode Islands 2009, S. 1-2.

Plattner, Zeier / In-Memory Data Management /

Hasso Plattner, Alexander Zeier: In-Memory Data Management. An Inflection Point for Enterprise Applications. Berlin - Heidelberg 2011.

Poggensee / Investitionsrechnung /

Kay Poggensee: Investitionsrechnung. Grundlagen – Aufgaben – Lösungen. Wiesbaden 2009.

Polli, Cook / Product Life Cycle /

Rolando Polli, Victor Cook: Validity of the Product Life Cycle. In: Journal of Business, Band 42, Nr. 4, 1969, S. 385-400.

Posner / International Trade /

Michael Posner: International Trade and Technical Change. In: Oxford Economic Papers, Band 13, Nr. 3, 1961, S. 323-341.

Poulton / Data Storage Networking /

Nigel Poulton: Data Storage Networking. Real-World Skills for the CompTIA Storage+ Certification and Beyond. Indiana 2014.

Preißler / Kosten- und Leistungsrechnung /

Peter Preißler: Entscheidungsorientierte Kosten- und Leistungsrechnung. 3. Auflage, München 2005.

Railsback, Grimm / Agent-Based Modeling /

Steven Railsback, Volker Grimm: Agent-Based and Individual-Based Modeling. A Practical Introduction. Princeton 2012.

Rau / Statistik /

Thomas Rau: Planung, Statistik und Entscheidung. Betriebswirtschaftliche Instrumente für die Kommunalverwaltung. München 2004.

Rautenstrauch, Schulze / Informatik /

Claus Rautenstrauch, Thomas Schulze: Informatik für Wirtschaftswissenschaftler und Wirtschaftsinformatiker. Berlin - Heidelberg 2003.

Reichman, Whiteley, Chi / File Storage Costs /

Andrew Reichman, Robert Whiteley, Eric Chi: File Storage Costs Less In The Cloud Than In-House. In: Forrester Report for Infrastructure & Operations Professionals. Cambridge 2011.

Reid, Fraser-King, Schwaderer / Managing Data /

Roger Reid, Gareth Fraser-King, David Schwaderer: Data Lifecycles. Managing Data for Strategic Advantage. Hoboken et al. 2007.

Reiner et al. / ILM EMC /

David Reiner, Gil Press, Mike Lenaghan, David Barta, Rich Urmston: Information Lifecycle Management: The EMC Perspective. In: Proceedings of the 20th International Conference on Data Engineering. Boston 2004, S. 804-808.

Repschläger, Wind, Zarnekow / Anbietersauswahl in der Cloud /

Jonas Repschläger, Stefan Wind, Rüdiger Zarnekow: Klassifikationsrahmen für die Anbietersauswahl in der Cloud. In: 41. Jahrestagung der Gesellschaft für Informatik: IOS 2.0 – Neue Aspekte der Zwischenbetrieblichen Integration durch Enterprise 2.0. Berlin 2011, S. 1-15.

Riesenhuber / Empirische Forschung /

Felix Riesenhuber: Großzahlige empirische Forschung. In: Sönke Albers, Daniel Klapper, Udo Konradt, Achim Walter, Joachim Wolf (Hrsg.): Methodik der empirischen Forschung. 3. Auflage, Wiesbaden 2009, S.1-18.

Rimal, Choi, Lumb / Cloud Ecosystems /

Bhaskar Rimal, Eunmi Choi, Ian Lumb: A Taxonomy, Survey and Issues of Cloud Computing Ecosystems. In: Nick Antonopoulos, Lee Gillam (Hrsg.): Cloud Computing. Principles, Systems and Applications, London et al. 2010, S. 21-46.

Rittinghouse, Ransome / Cloud Computing /

John Rittinghouse, James Ransome: Cloud Computing: Implementation, Management and Security. Boca Raton 2009.

Röhrich / Investitionsrechnung /

Martina Röhrich: Grundlagen der Investitionsrechnung. Eine Darstellung Anhand einer Fallstudie. München 2007.

Rolfes / Investitionsrechnung /

Bernd Rolfes: Moderne Investitionsrechnung. 3. Auflage, München 2003.

Rossi, Sein / Design Research /

Matti Rossi, Maung Sein: Design research workshop: A proactive research approach. In: 26th Information Systems Research Seminar in Scandinavia, Haikko 2003, S. 9-12.

Rüdiger / Stückwerk ILM /

Ariane Rüdiger: Stückwerk ILM. In: ECMguide-Magazin, Nr. 2, 2006, S. 18-19.

Rüter et al. / IT-Governance /

Andreas Rüter, Jürgen Schröder, Axel Göldner, Jens Niebuhr: IT-Governance in der Praxis. 2. Auflage, Heidelberg 2010.

Rumpel / IT-Infrastrukturen /

Rainer Rumpel: Planung und Realisierung von IT-Infrastrukturen – Ein prozess-basierter Ansatz. München 2012.

Sachs, Hedderich / Angewandte Statistik /

Lothar Sachs, Jürgen Hedderich: Angewandte Statistik. Methodensammlung mit R. 12. Auflage, Berlin - Heidelberg - New York 2007.

Satyanarayanan / Study Lifetimes /

Mahadev Satyanarayanan: A Study of File Sizes and Functional Lifetimes. In: Proceedings of the 8th ACM Symposium on Operating Systems Principles, 1981, S. 96-108.

Schadler / Email in the Cloud /

Ted Schadler: Should Your Email Live In The Cloud? A Comparative Cost Analysis. Cambridge 2009.

Scharer, Hohmann / Wirtschaftlichkeit IT /

Matthias Scharer, Joachim Hohmann: Wirtschaftlichkeit des Einsatzes von CAFM. In: Michael May (Hrsg.): CAFM-Handbuch. IT im Facility Management erfolgreich einsetzen. 3. Auflage, Berlin 2013, S. 113-146.

Schauer, Frank / WI und IS /

Carola Schauer, Ulrich Frank: Wirtschaftsinformatik und Information Systems - Ein Vergleich aus wissenschaftstheoretischer Sicht. In: Franz Lehner, Stephan Zelewski (Hrsg.): Wissenschaftstheoretische Fundierung und wissenschaftliche Orientierung der Wirtschaftsinformatik. Guito - Berlin 2007, 121-154.

Scheer / Geschäftsprozess /

August Wilhelm Scheer: ARIS – Vom Geschäftsprozess zum Anwendungssystem. 4. Auflage, Berlin - Heidelberg - New York 2002.

Scheffler / Steuerbilanz /

Wolfram Scheffler: Besteuerung von Unternehmen II. Steuerbilanz. 7. Auflage, Heidelberg et al. 2011.

Scheuermann / Simulationen /

Günter Scheuermann: Simulationen mit Inventor. FEM und dynamische Simulationen. Grundlagen und Beispiele ab Version 2013. München 2012.

Schierenbeck, Wöhle / Betriebswirtschaftslehre /

Henner Schierenbeck, Claudia Wöhle: Grundzüge der Betriebswirtschaftslehre. 18. Auflage, München 2012.

Schmeisser et al. / Finanzierung und Investition /

Wilhelm Schmeisser, Gerfried Hannemann, Dieter Krimphove, Marc Toebe, Horst Zündorf: Finanzierung und Investition. München 2012.

Schmidt / Kostenrechnung /

Andreas Schmidt: Kostenrechnung. Grundlagen der Vollkosten-, Deckungsbeitrags- und Plankostenrechnung sowie des Kostenmanagements. 5. Auflage, Stuttgart 2008.

Schmidt / Wirtschaftlichkeit /

Jürgen Schmidt: Wirtschaftlichkeit in der öffentlichen Verwaltung. 7. Auflage, Berlin 2006.

Schmidt, Klüver, Klüver / Soft Computing /

Jörn Schmidt, Christina Klüver, Jürgen Klüver: Programmierung naturanaloger Verfahren. Soft Computing und verwandte Methoden. Wiesbaden 2010.

Schmitz / Migrationsstrategie /

Carolin Schmitz: Entwicklung einer optimalen Migrationsstrategie für ein hierarchisches Datenmanagement System.

<http://www2.fz-juelich.de/jsc/docs/printable/ib/ib-04/ib-2004-07.pdf>, Jülich 2004,

Abruf: 2011-04-10.

Schnjakin, Meinel / Speicherressourcen in der Cloud /

Maxim Schnjakin, Christoph Meinel: Plattform zur Bereitstellung sicherer und hochverfügbarer Speicherressourcen in der Cloud. In: 12. Deutscher IT-Sicherheitskongress des BSI. Bonn 2011, S. 1-11.

Schoop / Informationsmanagement /

Schoop: Informationsmanagement. In: WISU Band 28, Nr. 4, 1999, S. 556-568.

Schröder / Operatives Controlling /

Harry Schröder: Operatives Controlling zielorientiert umsetzen. Der Leitfaden für eine erfolgreiche operative Controlling-Umsetzung. Düsseldorf 2012.

Schröder / Zeitreihenprognose /

Michael Schröder: Einführung in die kurzfristige Zeitreihenprognose und Vergleich der einzelnen Verfahren. In: Peter Mertens, Susanne Rässler (Hrsg.): Prognoserechnung. 6. Auflage, Heidelberg 2004, S. 7-37.

Schütte / Einsatzpotential In-Memory /

Reinhard Schütte: Analyse des Einsatzpotentials von In-Memory-Technologien in Handelsinformationssystemen. In: Wolfgang Lehner, Gunther Piller (Hrsg.): Innovative Unternehmensanwendungen mit In-Memory Data Management, IMDM 2011, Mainz 2011, S. 1-12.

Schütte / Modelle /

Reinhard Schütte: Grundsätze ordnungsgemäßer Referenzmodellierung. Konstruktion konfigurations- und anpassungsorientierter Modelle. Wiesbaden 1998.

Schwarze / Wirtschaftsinformatik /

Jochen Schwarze: Einführung in die Wirtschaftsinformatik. 4. Auflage, Herne 1997.

Shah et al. / Classification for ILM /

Gauri Shah, Kaladhar Voruganti, Piyush Shivam, Maria Alvarez: ACE: Classification for Information Lifecycle Management. Almaden et al. 2006.

Shelly, Rosenblatt / Systems Design /

Gary Shelly, Harry Rosenblatt: Systems Analysis and Design. 9. Auflage, Boston 2012.

Shen et al. / Data Protection in the Cloud /

Yushi Shen, Yale Li, Ling Wu, Shaofeng Liu, Qian Wen: Data Protection in the Cloud Era. In: Yushi Shen, Yale Li, Ling Wu, Shaofeng Liu, Qian Wen (Hrsg.): Enabling the New Era of Cloud Computing. Data Security, Transfer and Management. Hershey 2014, S. 132-153.

Shirazi / Performance Tuning /

Jack Shirazi: Java Performance Tuning. 2. Auflage, Sebastopol 2003.

Short / ILM /

James Short: Information Lifecycle Management: An Analysis of End User Perspectives. San Diego 2006.

Simon / Artificial /

Herbert Simon: The Sciences of the Artificial. 3. Auflage. Cambridge-Massachusetts 1996.

Skyrme / Add value to business /

David Skyrme: Ten Ways to Add Value to Your Business. In: Aslib Association for Information Management (Hrsg.): Managing Information, Band 1, Nr.3, 1994, S. 20-25.

Smullen et al. / Unstructured Data Processing Applications /

Clinton Smullen, Shahrukh Tarapore, Sudhanva Gurumurthi, Parthasarathy Ranganathan, Mustafa Uysal: Active Storage Revisited: The Case for Power and Performance Benefits for Unstructured Data Processing Applications. In: Proceedings of the 5th conference on Computing frontiers, Ischia 2008, S. 293-304.

Sneed, Baumgartner, Seidl / Systemtest /

Harry Sneed, Manfred Baumgartner, Richard Seidl: Der Systemtest. München 2009.

SNIA / ILM Definition /

Storage Networking Industry Association: ILM Definition and Scope. An ILM Framework.

http://www.snia.org/forums/dmf/knowledge/archives/SNIAs_Vision_for_ILM-SNWOct04.pdf, o. O. 2004,

Abruf: 2011-04-07.

SNIA / ILM Model /

Storage Networking Industry Association: The Information Lifecycle Management Maturity Model.

http://www.snia.org/sites/default/files/SNIA-DMF_ILM_Maturity_Model_20090921-Final.pdf, o. O 2009,

Abruf: 2014-07-21.

Sollbach, Thome / ILM Prozessimplementierung /

Wolfgang Sollbach, Günter Thome: Information Lifecycle Management. Prozessimplementierung. Berlin - Heidelberg 2008.

Sosinsky / Cloud Computing Bible /

Barrie Sosinsky: Cloud Computing Bible. Indianapolis 2011.

Sotomayor et al. / Private and Hybrid Clouds /

Borja Sotomayor, Ruben Montero, Ignacio Llorente, Ian Foster: Virtual Infrastructure Management in Private and Hybrid Clouds. In: IEEE Internet Computing, Band 13, Nr. 5, 2009, S. 14-22.

Spasojevic, Satyanarayanan / Study File System /

Mirjana Spasojevic, Mahadev Satyanarayanan: An Empirical Study of a WideArea Distributed File System. In: ACM Transactions on Computer Systems, Band 14, Nr. 2, 1996, S. 200-222.

Sriram / Cloud-Scale /

Ilango Sriram: SPECI, a Simulation Tool Exploring Cloud-Scale Data Centres. In: Martin Jaatun, Gansen Zhao, Chunming Rong (Hrsg.): Cloud Computing: Proceedings of the First International Conference, CloudCom 2009, Beijing 2009, S. 381-392.

Stahlknecht, Hasenkamp / Wirtschaftsinformatik /

Peter Stahlknecht, Ulrich Hasenkamp: Einführung in die Wirtschaftsinformatik. 11. Auflage, Berlin - Heidelberg - New York 2005, S. 135-137.

Stankov, Datsenka / Global Software Development /

Ivo Stankov, Rastsislan Datsenka: Platform-as-a-Service as an Enabler for Global Software Development and Delivery. In: Matthias Schumann, Lutz Kolbe, Michael Breitner, Arne Frerichs (Hrsg.): Multikonferenz Wirtschaftsinformatik 2010. Göttingen 2010, S. 113-116.

Starke / Software-Architekturen /

Gernot Starke: Effektive Software-Architekturen. München 2008.

Steinbinder, Bedau, Löw / Datenbanktechnologie für das ILM /

Detlev Steinbinder, Christoph Bedau, Jörg Löw: Nutzung spaltenbasierter Datenbanktechnologie für das Information Lifecycle Management in Anwendungssystemen. In: Thomas Barton, Burkhard Erdlenbruch, Frank Herrmann, Christian Müller (Hrsg.): Herausforderungen an die Wirtschaftsinformatik: Betriebliche Anwendungssysteme. Tagungsband zur AKWI-Fachtagung, Worms 2011, S. 101-114.

Steinhardt / ROI of Storage /

Ken Steinhardt: The ROI of Storage. Start with a Strategy. In: CSO, Band 6, Nummer 4, 2004, S. 6-7.

Stelling / Controlling /

Johannes Stelling: Kostenmanagement und Controlling. 2. Auflage, München 2005.

Stelzer / Sicherheitsstrategien /

Dirk Stelzer: Sicherheitsstrategien in der Informationsverarbeitung. Ein objekt-orientiertes, wissensbasiertes System für die Risikoanalyse. Wiesbaden 1993.

Stepan, Fischer / Optimierung /

Adolf Stepan, Edwin Fischer: Betriebswirtschaftliche Optimierung. Einführung in die quantitative Betriebswirtschaftslehre. 8. Auflage, München 2009.

Stötzer / Performance Reporting /

Sandra Stötzer: Stakeholder Performance Reporting von Nonprofit-Organisationen. Dissertation vorgelegt am Institut für Betriebswirtschaftslehre der Johannes Kepler Universität Linz, Linz 2009.

Stollenwerk / Wertschöpfungsmanagement /

Andreas Stollenwerk: Wertschöpfungsmanagement im Einkauf. Analysen - Strategien - Methoden - Kennzahlen. Wiesbaden 2012.

SUN / Storage Optimization /

SUN Microsystems: Storage Optimization. Information Lifecycle Management. Santa Clara 2006.

Sun, Kantor / Information System Evaluation /

Ying Sun, Paul Kantor: A new model for information system evaluation. In: Journal of the American Society for Information Science and Technology, Band 57, Nr. 5, 2006, S. 614-628.

Suter / Marktübersicht Cloud /

Valentin Suter: Die Cloud als Primärspeicher. Marktübersicht Cloud-Storage-Angebote. In: Swiss IT Magazine, Nr.3, 2012, S. 32-34.

Szyska / Kostenrechnung /

Uwe Szyska: Operatives Controlling auf Basis IT-gestützter Kostenrechnung. Wiesbaden 2011.

Talburdt / Entity Resolution /

John Talburdt: Entity Resolution and Information Quality. Burlington 2011.

Tauberger / Controlling

André Tauberger: Controlling für die öffentliche Verwaltung. München 2008.

Terplan, Voigt / Cloud Computing /

Kornel Terplan, Christian Voigt: Cloud Computing. Heidelberg et al. 2011.

Thome, Sollbach / Modelle des ILM /

Günter Thome, Wolfgang Sollbach: Grundlagen und Modelle des Information Lifecycle Management. Springer-Verlag, Berlin - Heidelberg - New York 2007.

Thonemann / Operations Management /

Ulrich Thonemann: Operations Management. Konzepte, Methoden und Anwendungen. 2. Auflage, München 2010.

Töpfer / Betriebswirtschaftslehre /

Armin Töpfer: Betriebswirtschaftslehre. Anwendungs- und prozessorientierte Grundlagen. 2. Auflage, Berlin - Heidelberg 2007.

Troppens et al. / Storage Networks /

Ulf Troppens, Wolfgang Müller-Friedt, Rainer Wolafka, Rainer Erkens, Nils Hausteint: Storage Networks Explained. Basics and Application of Fibre Channel SAN, NAS, iSCSI, Infiniband and FCoE. 2. Auflage, Heidelberg 2009.

Tsvihun, Stephanow, Streitberger / Sicherheit Cloud Computing /

Iryna Tsvihun, Philipp Stephanow, Werner Streitberger: Vergleich der Sicherheit traditioneller IT-Systeme und Public Cloud Computing Systeme. Arbeitspapier Fraunhofer Institut für Sichere Informationstechnologie SIT. München 2010.

Turczyk / ILM /

Lars Arne Turczyk: Information Lifecycle Management - Eine Methode zur Wertzuweisung von Dateien. Dissertation vorgelegt am Fachbereich Elektrotechnik und Informationstechnik der Technischen Universität Darmstadt, Darmstadt 2009.

Turczyk et al. / Model ILM /

Lars Arne Turczyk, Oliver Heckmann, Rainer Berbner, Ralf Steinmetz: An Analytical Model for Information Lifecycle Management. In: Mehdi-Khosrow Pour (Hrsg.): Emerging Trends and Challenges in Information Technology Management. Hershey 2006, S. 527-530.

Turczyk / Organisation ILM /

Lars Arne Turczyk: Information Lifecycle Management: Organisation ist wichtiger als Technologie. In: Information, Wissenschaft und Praxis, Nr. 7, 2005, S. 371-372.

Turczyk / Wertzuweisung /

Lars Arne Turczyk: Eine Methode zur Wertzuweisung von Dateien in ILM. In: Martin Bichler, Thomas Hess, Helmut Krcmar, Ulrike Lechner, Florian Matthes, Arnold Picot, Benjamin Speitkamp, Petra Wolf (Hrsg.): Multikonferenz Wirtschaftsinformatik 2008. Berlin 2008, S. 459-470.

UC Berkeley / SAN Service /

University of California, Berkeley: Storage Area Network (SAN) Services. <http://ist.berkeley.edu/is/platforms/storage/san/>; Berkeley 2014, Abruf: 2014-10-02.

Urbanski / Cloud-Outcomes /

Jürgen Urbanski: IT at a Crossroads: Driving the Transformation of IT-Services Production Towards Cloud-Enabled Business Outcomes. In: Hagen Rickmann, Stefan Diefenbach, Kai Brüning (Hrsg.): IT-Outsourcing. Neue Herausforderungen im Zeitalter von Cloud-Computing, Berlin - Heidelberg 2013, S. 39-70.

Vahrenkamp / Produktionsmanagement /

Richard Vahrenkamp: Produktionsmanagement. 6. Auflage, München 2008.

van Husen / IT-Dienstleistungen /

Christian van Husen: IT-Service Management Konfigurator. In: Klaus-Peter Fähnrich, Christian van Husen (Hrsg.): Entwicklung IT-basierter Dienstleistungen, Heidelberg 2008, S. 61-68.

Vassiliadis et al. / Warehouse Process Management /

Panos Vassiliadis, Christoph Quix, Yannis Vassiliou, Matthias Jarke: Data Warehouse Process Management. In: Journal of Information Systems, Band 26, Nr.3, 2001, S. 205-236.

VDI / Betriebswirtschaftliche Berechnungen /

Verein Deutscher Ingenieure: Betriebswirtschaftliche Berechnungen für Investitionsgüter und Anlagen. VDI-Richtlinie 6025, Düsseldorf 2012.

VDI / Simulation /

Verein Deutscher Ingenieure: Simulation von Logistik-, Materialfluss- und Produktionssystemen. VDI-Richtlinie 3633, Düsseldorf 2013.

Velten / Modeling and Simulation /

Kai Velten: Modeling and Simulation. Introduction for Scientists and Engineers. Weinheim 2009.

Venable / Design Science Research /

John Venable: Design Science Research Post Hevner et al.: Criteria, Standards, Guidelines and Expectations. In: Robert Winter, Leon Zhao, Stephan Aier (Hrsg.): Global Perspectives on Design Science Research. DESRIST'10 Proceedings of the 5th international conference on Design Science Research in Information Systems, Band 6105, 2010, S. 109-123.

Venable / Framework for Design Science Research Activities /

John Venable: A Framework for Design Science Research Activities. In: Mehdi Khosrow-Pour (Hrsg.): Emerging Trends and Challenges in Information Technology Management. Proceedings of the 2006 Information Resource Management Association Conference, Band 1, S. 184-187.

Venable, Pries-Heje, Baskerville / Framework for Evaluation /

John Venable, Jan Pries-Heje, Richard Baskerville: A Comprehensive Framework for Evaluation in Design Science Research. In: Ken Peffers, Marcus Rothenberger, Bill Kuechler (Hrsg.): Design Science Research in Information Systems. Advances in Theory and Practice. DESRIST'12 Proceedings of the 7th international conference on Design Science Research in Information Systems, Band 7286, 2012, S. 423-438.

Venhaus, Haselbeck, Wintermann / Schatten der Überwachung /

Marc Venhaus, Sebastian Haselbeck, Ole Wintermann: Globalisierung im Schatten der Überwachung. Internet. Demokratie. Handel. 1. Auflage, Berlin 2013.

Ventzislavova, Hensel / Betriebswirtschaftliche Formelsammlung /

Madlen Ventzislavova, Christian Hensel: Betriebswirtschaftliche Formelsammlung. Finanzierung, Investition, Rechnungswesen, Kosten-Leistungsrechnung, Steuern. 4. Auflage, Norderstedt 2012.

Verma et al. / Very Large File Systems /

Akshat Verma, Upendra Sharma, Jim Rubas, David Pease, Marc Kaplan, Rohit Jain, Murthy Devarakonda, Mandis Beigi: An Architecture for Lifecycle Management in Very Large File Systems. In: 13th NASA Goddard, 22nd IEEE Conference on Mass Storage Systems and Technologies (MSST 2005), Monterey 2005, S. 160-168.

Viereck, Sonderhüsken / Informationstechnik in der Praxis /

Axel Viereck, Bernhard Sonderhüsken: Informationstechnik in der Praxis. Eine Einführung in die Wirtschaftsinformatik. Stuttgart - Leipzig - Wiesbaden 2001.

Voegele, Sommer / Wirtschaftlichkeitsrechnung /

Arno Voegele, Lutz Sommer: Kosten- und Wirtschaftlichkeitsrechnung für Ingenieure. Kostenmanagement im Engineering. München 2012.

vom Brocke, Buddendick / Konstruktionstechniken Referenzmodellierung /

Jan vom Brocke, Christian Buddendick: Konstruktionstechniken für die Referenzmodellierung - Systematisierung, Sprachgestaltung und Werkzeugunterstützung. In: Jörg Becker, Patrick Delfmann (Hrsg.): Referenzmodellierung. Grundlagen, Techniken und domänenbezogene Anwendungen. Berlin - Heidelberg 2004, S. 19-50.

Vouk / Cloud Computing /

Mladen Vouk: Cloud Computing: Issues, Research and Implementations. In: Journal of Computing and Information Technology, Band 16, Nr. 4, 2008, S. 235-246.

Wallmüller / Software-Qualitätsmanagement /

Ernest Wallmüller: Software-Qualitätsmanagement in der Praxis. 2. Auflage, München - Wien 2001.

Walls, Widmeyer, El Sawy / Design Theory /

Joseph Walls, George Widmeyer, Omar El Sawy: Building an information system design theory for vigilant EIS. In: Information Systems Research, Band 3, Nr. 1, 1992, S. 36-59.

Wang et al. / Cloud Methodology /

Lizhe Wang, Rajiv Ranjan, Jinjun Chen, Boualem Benatallah: Cloud Computing Methodology, Systems and Applications. 1. Auflage, Boca Raton 2012.

Webster, Watson / Literature Review /

Jane Webster, Richard Watson: Analyzing the past to prepare for the future: Writing a literature review. In: MIS Quarterly, Band 26, Nr. 2, 2002, S. 13-23.

Weinhardt et al. / Cloud-Computing / 460

Christof Weinhardt, Arun Anandasivam, Benjamin Blau, Nikolay Borissov, Thomas Meinl, Wibke Michalk, Jochen Stößer: Cloud-Computing: Eine Abgrenzung, Geschäftsmodelle und Forschungsgebiete. In: Wirtschaftsinformatik, Band 51, Nr. 5, 2007, S. 453-462.

Weiss / Computing in the Clouds /

Aaron Weiss: Computing in the Clouds. In: networker – Cloud Computing: PC functions move onto the web, Band 11, Nr. 4, 2007, S. 16-25.

Wicker / Death of Privacy /

Stephen Wicker: Cellular Convergence and the Death of Privacy. 1. Auflage, New York 2013.

Wieczorrek, Mertens / Management IT-Projekte /

Wilhelm Wieczorrek, Peter Mertens: Management von IT-Projekten. 3. Auflage, Berlin - Heidelberg 2008.

Wijnhoven, Amrit / Use-based file retention method /

Fons Wijnhoven, Chintan Amrit: Evaluating the Applicability of a Use Value-Based File Retention Method. In: Proceedings of SIGSVC Workshop, o.O. 2010, S. 1-14

Winkler / Securing the Cloud /

Vic Winkler: Securing the Cloud. Cloud Computer Security Techniques and Tactics. 1. Auflage, Waltham 2011.

Witten et al. / Documents and Metadata /

Ian Witten, David Bainbridge, Gordon Paynter, Stefan Boddie: Importing Documents and Metadata into Digital Libraries: Requirements Analysis and Extensible Architecture. In: Maristella Agosti, Constantino Thanos (Hrsg.): Proceedings of the 6th European Conference of Research and Advanced Technology for Digital Libraries, Rom 2002, S. 389-405.

Wittenburg, Trilsbeek, Wittenburg / Archiving and Data Management /

Peter Wittenburg, Paul Trilsbeek, Florian Wittenburg: Corpus Archiving and Dissemination. In: Jacques Durand, Ulrike Gut, Gjert Kristoffersen (Hrsg.): The Oxford Handbook of Corpus Phonology, Oxford 2014, S. 133-149.

Wöhe, Döring / Betriebswirtschaftslehre /

Günter Wöhe, Ulrich Döring: Einführung in die Allgemeine Betriebswirtschaftslehre. 23. Auflage, München 2008.

Wöltje / Investition und Finanzierung /

Jörg Wöltje: Investition und Finanzierung. Grundlagen, Verfahren, Übungsaufgaben und Lösungen. Freiburg 2013.

Wolff / IT-Controlling /

Frank Wolff: Ökonomie multiperspektivischer Unternehmensmodellierung. IT-Controlling für modell-basiertes Wissensmanagement. Dissertation vorgelegt an der Fakultät für Wirtschaftswissenschaften der Universität Duisburg-Essen, Wiesbaden 2008.

Wootton / Developing Quality Metadata /

Cliff Wootton: Developing Quality Metadata. Building Innovative Tools and Workflow Solutions. 1. Auflage, Burlington 2007.

Wünsche / Finanzwirtschaft /

Manfred Wünsche: Finanzwirtschaft der Bilanzbuchhalter. 3. Auflage, Wiesbaden 2010.

Yates-Mercer, Bawden / Managing the paradox /

Penny Yates-Mercer, David Bawden: Managing the paradox: the valuation of knowledge and knowledge management. In: Journal of Information Science, Band 28, Nr.1, 2002, S. 19-29.

Yuan, Yang, Chen / Storage in the Cloud /

Dong Yuan, Yun Yang, Jinjun Chen: Computation and Storage in the Cloud. Understanding the Trade-Offs. 1. Auflage, London 2013.

Zadok et al. / Reducing Storage Management Costs /

Erez Zadok, Jeffrey Osborn, Ariye Shater, Charles Wright, Kiran-Kumar Muniswamy-Reddy, Jason Nieh: Reducing Storage Management Costs via Informed User-Based Policies. In: 12th NASA Goddard, 21st IEEE Conference on Mass Storage Systems and Technologies (MSST 2004), Maryland 2004, S. 101-105.

Zangemeister / Nutzwertanalyse /

Christof Zangemeister: Nutzwertanalyse in der Systemtechnik. 4. Auflage, München 1976.

Zantow, Dinauer / Finanzwirtschaft /

Roger Zantow, Josef Dinauer: Finanzwirtschaft des Unternehmens: Die Grundlagen des modernen Finanzmanagements. 3. Auflage, München 2011.

Zarnekow / IT-Dienstleistungen /

Rüdiger Zarnekow: Produktionsmanagement von IT-Dienstleistungen. Grundlagen, Aufgaben und Prozesse. Berlin - Heidelberg - New York 2007.

Zarnekow, Brenner / Kosten IT /

Rüdiger Zarnekow, Walter Brenner: Einmalige und wiederkehrende Kosten von IT-Anwendungen: eine empirische Untersuchung. In: Zeitschrift für Controlling und Management. Band 48, Nr. 5, 2004, S. 336-339.

Zelewski / Publikationspraxis /

Stephan Zelewski: Kann Wissenschaftstheorie behilflich für die Publikationspraxis sein? In: Franz Lehner, Stephan Zelewski (Hrsg.): Wissenschaftstheoretische Fundierung und wissenschaftliche Orientierung der Wirtschaftsinformatik. Berlin 2007, S. 71-120.

Zelewski, Hohmann, Hügens / Planungs- und Steuerungssysteme /

Stephan Zelewski, Susanne Hohmann, Torben Hügens: Produktionsplanungs- und -steuerungssysteme. Konzepte und exemplarische Implementierungen mithilfe von SAP R/3. München 2008.

Zimmermann, Fries, Hoch / Rechnungswesen /

Werner Zimmermann, Hans-Peter Fries, Gero Hoch: Betriebliches Rechnungswesen. 8. Auflage, München 2003.

Zimmermann, Henke, Broer / Finanzwissenschaft /

Horst Zimmermann, Klaus-Dirk Henke, Michael Broer: Finanzwissenschaft. 10. Auflage, München 2011.

Zimmermann, Stache / Quantitative Methoden /

Werner Zimmermann, Ulrich Stache: Operations Research. Quantitative Methoden zur Entscheidungsvorbereitung. 10. Auflage, München 2001.

Zsifkovits / Projektsteuerung /

Helmut Zsifkovits: Statusüberwachung und Projektsteuerung. In: Ernst Tiemeyer (Hrsg.): Handbuch IT-Projektmanagement. 2. Auflage, München 2014.

Anhang 1 Datenmengen und Wachstum der Daten

In Abbildung A1-1 ist dargestellt, wie sich die Datenmengen, ausgehend von 1 Terabyte, 10 Terabyte und 100 Terabyte im ersten Jahr, bei einem jährlichen Wachstum von 30 Prozent in den Jahren zwei, drei, vier und fünf, vergrößern. Diese Datenmengen bilden die Grundlage zur Berechnung der Speicher- und Verlagerungskosten¹²⁰⁶ im Rahmen der betrachteten Szenarien.¹²⁰⁷

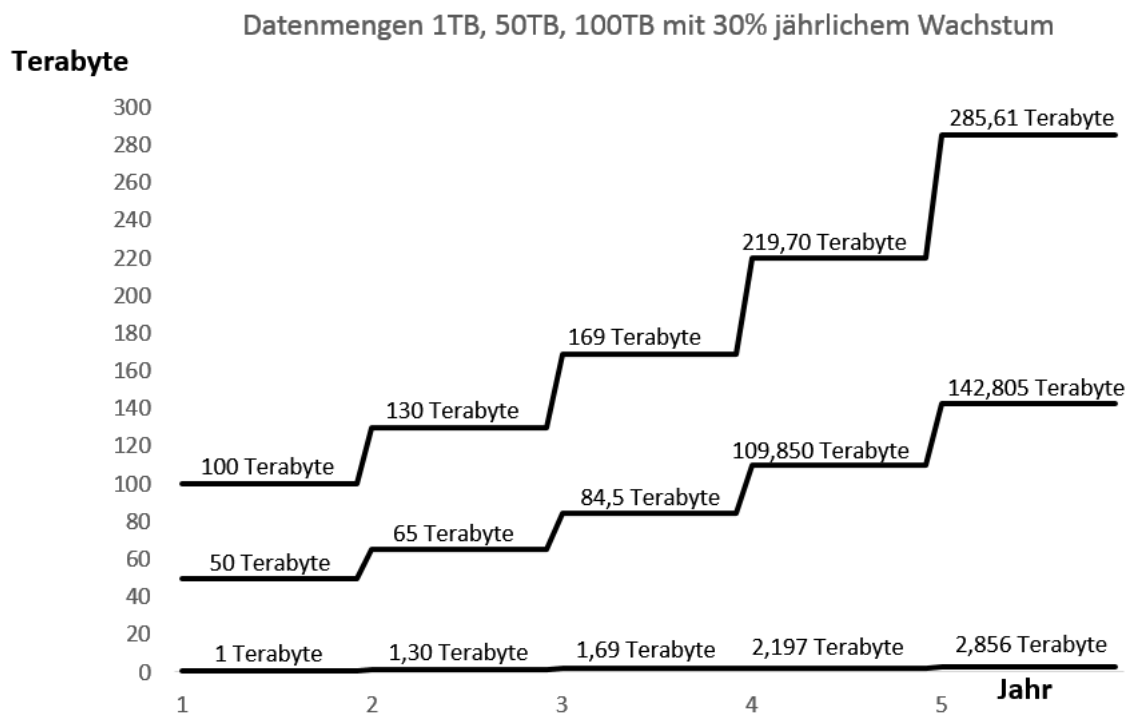


Abb. A1-1: Datenmengen und Wachstum der Daten

¹²⁰⁶ Vgl. Kapitel 5.3

¹²⁰⁷ Vgl. Kapitel 5.3

Anhang 2 Nutzungsgrade ohne Rauschfaktor

Sinkender Nutzungsgrad (ohne Rauschfaktor)							Steigender Nutzungsgrad (ohne Rauschfaktor)							Konstanter Nutzungsgrad (ohne Rauschfaktor)						
Monat	1	2	3	4	5	6	Monat	1	2	3	4	5	6	Monat	1	2	3	4	5	6
Jahr	1	1	1	1	1	1	Jahr	1	1	1	1	1	1	Jahr	1	1	1	1	1	1
Zugriffe	12.000	11.100	10.200	9.300	8.400	7.500	Zugriffe	3.000	2.000	1.000	500	0	0	Zugriffe	5.000	5.000	5.000	5.000	5.000	5.000
Monat	7	8	9	10	11	12	Monat	7	8	9	10	11	12	Monat	7	8	9	10	11	12
Jahr	1	1	1	1	1	1	Jahr	1	1	1	1	1	1	Jahr	1	1	1	1	1	1
Zugriffe	6.600	5.700	4.800	3.900	3.000	2.100	Zugriffe	0	0	0	0	0	0	Zugriffe	5.000	5.000	5.000	5.000	5.000	5.000
Monat	13	14	15	16	17	18	Monat	13	14	15	16	17	18	Monat	13	14	15	16	17	18
Jahr	2	2	2	2	2	2	Jahr	2	2	2	2	2	2	Jahr	2	2	2	2	2	2
Zugriffe	1.200	300	0	0	0	0	Zugriffe	300	600	900	1.200	1.500	1.800	Zugriffe	5.000	5.000	5.000	5.000	5.000	5.000
Monat	19	20	21	22	23	24	Monat	19	20	21	22	23	24	Monat	19	20	21	22	23	24
Jahr	2	2	2	2	2	2	Jahr	2	2	2	2	2	2	Jahr	2	2	2	2	2	2
Zugriffe	0	0	0	0	0	0	Zugriffe	2.100	2.400	2.700	3.000	3.300	3.600	Zugriffe	5.000	5.000	5.000	5.000	5.000	5.000
Monat	25	26	27	28	29	30	Monat	25	26	27	28	29	30	Monat	25	26	27	28	29	30
Jahr	3	3	3	3	3	3	Jahr	3	3	3	3	3	3	Jahr	3	3	3	3	3	3
Zugriffe	0	0	0	0	0	0	Zugriffe	3.900	4.200	4.500	4.800	5.100	5.400	Zugriffe	5.000	5.000	5.000	5.000	5.000	5.000
Monat	31	32	33	34	35	36	Monat	31	32	33	34	35	36	Monat	31	32	33	34	35	36
Jahr	3	3	3	3	3	3	Jahr	3	3	3	3	3	3	Jahr	3	3	3	3	3	3
Zugriffe	0	0	0	0	0	0	Zugriffe	5.700	6.000	6.300	6.600	6.900	7.200	Zugriffe	5.000	5.000	5.000	5.000	5.000	5.000
Monat	37	38	39	40	41	42	Monat	37	38	39	40	41	42	Monat	37	38	39	40	41	42
Jahr	4	4	4	4	4	4	Jahr	4	4	4	4	4	4	Jahr	4	4	4	4	4	4
Zugriffe	0	0	0	0	0	0	Zugriffe	7.500	7.800	8.100	8.400	8.700	9.000	Zugriffe	5.000	5.000	5.000	5.000	5.000	5.000
Monat	43	44	45	46	47	48	Monat	43	44	45	46	47	48	Monat	43	44	45	46	47	48
Jahr	4	4	4	4	4	4	Jahr	4	4	4	4	4	4	Jahr	4	4	4	4	4	4
Zugriffe	0	0	0	0	0	0	Zugriffe	9.300	9.600	9.900	10.200	10.500	10.800	Zugriffe	5.000	5.000	5.000	5.000	5.000	5.000
Monat	49	50	51	52	53	54	Monat	49	50	51	52	53	54	Monat	49	50	51	52	53	54
Jahr	5	5	5	5	5	5	Jahr	5	5	5	5	5	5	Jahr	5	5	5	5	5	5
Zugriffe	0	0	0	0	0	0	Zugriffe	11.100	11.400	11.700	12.000	12.300	12.600	Zugriffe	5.000	5.000	5.000	5.000	5.000	5.000
Monat	55	56	57	58	59	60	Monat	55	56	57	58	59	60	Monat	55	56	57	58	59	60
Jahr	5	5	5	5	5	5	Jahr	5	5	5	5	5	5	Jahr	5	5	5	5	5	5
Zugriffe	0	0	0	0	0	0	Zugriffe	12.900	13.200	13.500	13.800	14.100	14.400	Zugriffe	5.000	5.000	5.000	5.000	5.000	5.000

Abb. A2-1: Sinkender, steigender und konstanter Nutzungsgrad (ohne Rauschfaktor)¹²⁰⁸

Saisonaler Nutzungsgrad (ohne Rauschfaktor)							Sinkender Trend-Saisonaler Nutzungsgrad (ohne Rauschfaktor)							Steigender Trend-Saisonaler Nutzungsgrad (ohne Rauschfaktor)						
Monat	1	2	3	4	5	6	Monat	1	2	3	4	5	6	Monat	1	2	3	4	5	6
Jahr	1	1	1	1	1	1	Jahr	1	1	1	1	1	1	Jahr	1	1	1	1	1	1
Zugriffe	0	500	4.000	12.000	12.400	12.500	Zugriffe	10.500	10.800	11.300	12.000	12.400	12.500	Zugriffe	500	800	1.500	4.000	4.400	4.500
Monat	7	8	9	10	11	12	Monat	7	8	9	10	11	12	Monat	7	8	9	10	11	12
Jahr	1	1	1	1	1	1	Jahr	1	1	1	1	1	1	Jahr	1	1	1	1	1	1
Zugriffe	12.500	12.400	12.000	4.000	500	0	Zugriffe	12.500	12.400	12.000	9.500	8.800	8.500	Zugriffe	4.500	4.400	4.000	3.300	2.800	2.500
Monat	13	14	15	16	17	18	Monat	13	14	15	16	17	18	Monat	13	14	15	16	17	18
Jahr	2	2	2	2	2	2	Jahr	2	2	2	2	2	2	Jahr	2	2	2	2	2	2
Zugriffe	0	500	4.000	12.000	12.400	12.500	Zugriffe	8.500	8.800	9.300	10.000	10.400	10.500	Zugriffe	2.500	2.800	3.500	6.000	6.400	6.500
Monat	19	20	21	22	23	24	Monat	19	20	21	22	23	24	Monat	19	20	21	22	23	24
Jahr	2	2	2	2	2	2	Jahr	2	2	2	2	2	2	Jahr	2	2	2	2	2	2
Zugriffe	12.500	12.400	12.000	4.000	500	0	Zugriffe	10.500	10.400	10.000	7.500	6.800	6.500	Zugriffe	6.500	6.400	6.000	5.300	4.800	4.500
Monat	25	26	27	28	29	30	Monat	25	26	27	28	29	30	Monat	25	26	27	28	29	30
Jahr	3	3	3	3	3	3	Jahr	3	3	3	3	3	3	Jahr	3	3	3	3	3	3
Zugriffe	0	500	4.000	12.000	12.400	12.500	Zugriffe	6.500	6.800	7.300	8.000	8.400	8.500	Zugriffe	4.500	4.800	5.500	8.000	8.400	8.500
Monat	31	32	33	34	35	36	Monat	31	32	33	34	35	36	Monat	31	32	33	34	35	36
Jahr	3	3	3	3	3	3	Jahr	3	3	3	3	3	3	Jahr	3	3	3	3	3	3
Zugriffe	12.500	12.400	12.000	4.000	500	0	Zugriffe	8.500	8.400	8.000	5.500	4.800	4.500	Zugriffe	8.500	8.400	8.000	7.300	6.800	6.500
Monat	37	38	39	40	41	42	Monat	37	38	39	40	41	42	Monat	37	38	39	40	41	42
Jahr	4	4	4	4	4	4	Jahr	4	4	4	4	4	4	Jahr	4	4	4	4	4	4
Zugriffe	0	500	4.000	12.000	12.400	12.500	Zugriffe	4.500	4.800	5.300	6.000	6.400	6.500	Zugriffe	6.500	6.800	7.500	10.000	10.400	10.500
Monat	43	44	45	46	47	48	Monat	43	44	45	46	47	48	Monat	43	44	45	46	47	48
Jahr	4	4	4	4	4	4	Jahr	4	4	4	4	4	4	Jahr	4	4	4	4	4	4
Zugriffe	12.500	12.400	12.000	4.000	500	0	Zugriffe	6.500	6.400	6.000	3.500	2.800	2.500	Zugriffe	10.500	10.400	10.000	9.300	8.800	8.500
Monat	49	50	51	52	53	54	Monat	49	50	51	52	53	54	Monat	49	50	51	52	53	54
Jahr	5	5	5	5	5	5	Jahr	5	5	5	5	5	5	Jahr	5	5	5	5	5	5
Zugriffe	0	500	4.000	12.000	12.400	12.500	Zugriffe	2.500	2.800	3.300	4.000	4.400	4.500	Zugriffe	8.500	8.800	9.500	12.000	12.400	12.500
Monat	55	56	57	58	59	60	Monat	55	56	57	58	59	60	Monat	55	56	57	58	59	60
Jahr	5	5	5	5	5	5	Jahr	5	5	5	5	5	5	Jahr	5	5	5	5	5	5
Zugriffe	12.500	12.400	12.000	4.000	500	0	Zugriffe	4.500	4.400	4.000	1.500	800	500	Zugriffe	12.500	12.400	12.000	11.300	10.800	10.500

Abb. A2-2: Saisonaler, sinkender trend-saisonaler und steigender trend-saisonaler Nutzungsgrad (ohne Rauschfaktor)¹²⁰⁹

¹²⁰⁸ Vgl. Kapitel 5.2.2.1.3.1.1 bis 5.2.2.1.3.1.3

¹²⁰⁹ Vgl. Kapitel 5.2.2.1.3.1.4 bis 5.2.2.1.3.1.6

Anhang 3 Flussdiagramme für Klassifizierungsfunktionen

Anhang 3.1 Klassifizierungsfunktion K_1 für Szenario 1

Im ersten untersuchten Szenario¹²¹⁰ existieren drei unterschiedliche Klassen. Hierbei handelt es sich um die Klassen „Online“, „Nearline“ und „Offline“. Die Funktion $K_1(n, i, k)$ prüft, ob der Nutzungsgrad (n) innerhalb von zwei definierten Intervallen (i) enthalten ist.¹²¹¹ Wenn nicht, dann werden die Daten der Klasse des dritten Intervalls zugeordnet.

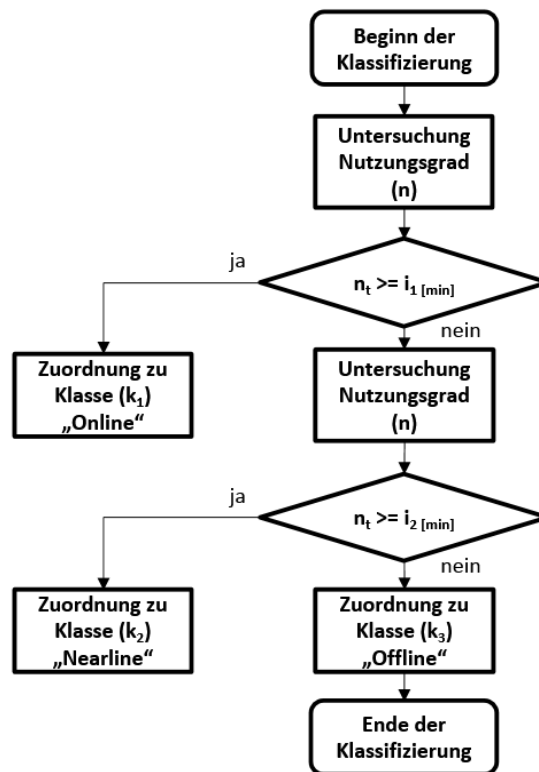


Abb. A3-1: Flussdiagramm für Klassifizierungsfunktion 1

¹²¹⁰ Vgl. Kapitel 5.2.3.1.1.1.1, Kapitel 5.3

¹²¹¹ Vgl. Kapitel 5.2.3.1

Anhang 3.2 Klassifizierungsfunktion K_2 für Szenario 2

Im zweiten Szenario¹²¹² existieren vier unterschiedliche Klassen. Hierbei handelt es sich um die Klassen „Cloud“, „Online“, „Nearline“ und „Offline“. Die Funktion $K_2(n, i, k)$ prüft, ob der Nutzungsgrad (n) innerhalb von drei definierten Intervallen (i) enthalten ist.¹²¹³ Wenn nicht, dann werden die Daten der Klasse des vierten Intervalls zugeordnet.

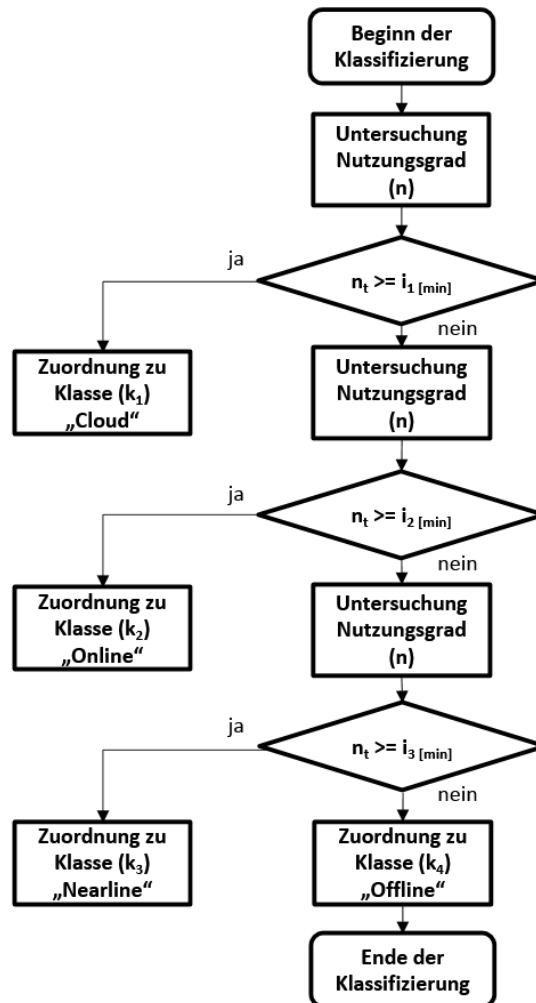


Abb. A3-2: Flussdiagramm für Klassifizierungsfunktion 2

¹²¹² Vgl. Kapitel 5.2.3.1.1.1.2

¹²¹³ Vgl. Kapitel 5.2.3.1

Q2

:

</

Abb. A4-2: Nutzungsgrade und Klassifizierungsfunktion Szenario 1

Anhang 4.2 Implementierung für Szenario 2

Die Tabelle „Parameter“ in Abbildung A4-3 verdeutlicht, wie die jeweiligen Klassen sowie Intervalle in Microsoft Excel 2013 für Szenario 2¹²¹⁷ implementiert wurden. Für die Klassen „4“, „3“, „2“ und „1“ bzw. „Cloud“, „Online“, „Nearline“ und „Offline“ werden die Intervalle zur Festlegung der Minimal- und Maximalwerte definiert.¹²¹⁸

A3

:

✕

✓

f_x

Klassen und Intervalle

	A	B	C	D
3	<u>Klassen und Intervalle</u>			
4	Klassen (k)	Intervalle (i)		
5	Wert der Klasse	Minimalwert [min]	Maximalwert [max]	Name der Klasse
6	4	8.000	∞	Cloud
7	3	3.000	7.999	Online
8	2	1	2.999	Nearline
9	1	0	0	Offline
	Parameter	Sink. Lebenszykl. (ohne Rausch)	Tabelle2	Tabelle3
			Tabelle4	Tabelle5

Abb. A4-3: Klassen und Intervalle Szenario 2

In Abbildung A4-4 sind beispielhaft Werte für den Monat, das Jahr und die Zugriffe für den Nutzungsgrad mit sinkenden Zugriffen dargestellt. Im Formelbereich ist die Klassifizierungsfunktion K₂ dargestellt. Diese prüft, ob die Zugriffe bzw. Nutzungsgrade in-

¹²¹⁷ Vgl. Kapitel 5.2.3.1.1; Für Szenario 2 wurde zusätzlich die Klasse „Cloud“ sowie das zugehörige Intervall implementiert.

¹²¹⁸ Vgl. Kapitel 5.2.3.1.1.2

nerhalb der Intervalle der Tabelle „Parameter“ enthalten sind und welchen Klassen die Daten zugeordnet werden müssen.




Q2	:	  	=(WENN(O2>=Parameter!\$B\$6; Parameter!\$A\$6; (WENN(O2>=Parameter!\$B\$7; Parameter!\$A\$7; (WENN(O2>=Parameter!\$B\$8; Parameter!\$A\$8; Parameter!\$A\$9))))))						
	M	N	O	Q	T	U	V	W	X
1	Monat	Jahr	Zugriffe (n)	Speicherebene					
2	1	1	12000	4					
3	2	1	11100	4					
4	3	1	10200	4					
5	4	1	9300	4					
6	5	1	8400	4					
7	6	1	7500	3					
8	7	1	6600	3					
9	8	1	5700	3					
10	9	1	4800	3					
Parameter			Sink. Lebenszykl. (ohne Rausch)	Tabelle2	Tabelle3	Tabelle4	Tabelle5	Tabelle6	

Abb. A4-4: Nutzungsgrade und Klassifizierungsfunktion Szenario 2

Anhang 5 Kostenarten und Kostenmodell

Anhang 5.1 Anschaffungskosten

Im Zuge der Einführung von ILM und der automatisierten Klassifizierung unstrukturierter Daten wird angenommen, dass die Mitarbeiter geschult werden müssen. Administratoren und IT-Verantwortliche erhalten eine Schulung zu bspw. den Inhalten Konzept des ILM¹²¹⁹, Vorgehensweise bei der Vorklassifizierung der unstrukturierten Daten, Auswahl von Werkzeugen zur Untersuchung der Metadaten, Verwaltung der Daten in unterschiedlichen Speicherebenen und zu den rechtlichen Anforderungen an die Verwaltung der Daten.¹²²⁰ Die für das erste Szenario angenommene Schulungsdauer beträgt 10 Tage. Bei Durchführung der Schulung durch einen externen Berater und einem angenommenen Tagessatz¹²²¹ von 1.000€ entstehen Schulungskosten von 10.000€. Für die Entwicklung und Implementierung eines ILM-Konzepts zur Verwaltung der Daten auf unterschiedlichen Speicherebenen und der Implementierung der Funktionen zur automatisierten Klassifizierung unstrukturierter Daten entstehen sonstige Anschaffungskosten. Hierfür werden Kosten in Höhe von 50.000€ angenommen, wenn die Entwicklung und

¹²¹⁹ Vgl. Kapitel 2

¹²²⁰ Vgl. Short / ILM / 21-22.

¹²²¹ Vgl. Lippold / Unternehmensberatung /194

Implementierung von externen Beratern¹²²² durchgeführt wird. Der angenommene Projektzeitraum beträgt 50 Tage.¹²²³ Für das zweite Szenario werden höhere Anschaffungskosten angenommen, da das ILM-Konzept um das Cloud Computing erweitert wird.¹²²⁴ Die Schulungskosten betragen 15.000€¹²²⁵ und die sonstigen Anschaffungskosten 60.000€^{1226, 1227}

Anhang 5.2 Betriebskosten - Parameter der Kostenfunktionen

In Tabelle A5-1 sind die Parameter aufgelistet und beschrieben, die in den Kostenfunktionen verwendet wurden.¹²²⁸

Parameter	Kostenart	Beschreibung
SBK _{Cloud; t}	Betriebskosten Speicherplatz (SBK)	Speicherplatzkosten Cloud-Ebene (je Gigabyte im Monat)
SBK _{Online; t}	Betriebskosten Speicherplatz (SBK)	Speicherplatzkosten Online-Ebene (je Gigabyte im Monat)
SBK _{Nearline; t}	Betriebskosten Speicherplatz (SBK)	Speicherplatzkosten Nearline-Ebene (je Gigabyte im Monat)
SBK _{Offline; t}	Betriebskosten Speicherplatz (SBK)	Speicherplatzkosten Offline-Ebene (je Gigabyte im Monat)
ABK _Z	Betriebskosten Administration (ABK)	Kosten für Zugriffe auf Daten in Cloud-Ebene (je Zugriff n _i) ¹²²⁹
BKF _{S; t}	Betriebskosten Administration (ABK) Allgemeine Betriebskosten (GBK)	Betriebskostenfaktor für Speichermedien
BKF _{C; t}	Betriebskosten Administration (ABK) Allgemeine Betriebskosten (GBK)	Betriebskostenfaktor für Cloud-Speicher
BKF _{V; t}	Betriebskosten Administration (ABK) Allgemeine Betriebskosten (GBK)	Betriebskostenfaktor für Verlagerung der Daten zwischen Speichermedien
V(K _{n; t})	Betriebskosten Verlagerung (VBK)	Verlagerungswert für erstes Szenario ¹²³⁰

¹²²² Der Tagessatz beträgt ebenfalls 1.000€.

¹²²³ Die Annahmen zum Projektzeitraum basieren auf einer Fallstudie zur Verwaltung strukturierter Daten im ILM. Vgl. hierzu Matthesius / ILM-Konzept Energiekonzern / 67-68.

¹²²⁴ Vgl. Kapitel 5.2.3.1.1.2

¹²²⁵ Bei einer angenommenen Schulungsdauer von 15 Tagen.

¹²²⁶ Bei einer angenommenen Projektdauer von 60 Tagen.

¹²²⁷ Der Anschaffungsaufwand bezüglich der Einführung von Cloud Computing ist unternehmensspezifisch. Er hängt von den Unternehmensprozessen und der Unternehmensorganisation ab. Vgl. hierzu Höllwarth / Cloud Migration / 184. In dieser Arbeit werden, im Vergleich zu den Aufwänden zum ILM-Konzept ohne Cloud Computing, für die Einführung und Implementierung von IaaS bzw. zur Nutzung von Cloud-Speicher einer Public Cloud, ein zusätzlicher Schulungsaufwand von 5 Tagen und ein zusätzlicher Projektaufwand von 10 Tagen angenommen.

¹²²⁸ Vgl. Kapitel 5.2.3.3.2

¹²²⁹ Amazon berechnet die Kosten für die Zugriffe auf die Daten der Cloud-Ebene je 10.000 Zugriffe. Vgl. hierzu Kapitel 5.2.3.3.2.

¹²³⁰ Vgl. Kapitel 5.2.3.2

Parameter	Kostenart	Beschreibung
$V_{CI}(K_n; t)$	Betriebskosten Verlagerung (VBK)	Verlagerungswert Cloud-Import für zweites Szenario ¹²³¹
$V_{CE}(K_n; t)$	Betriebskosten Verlagerung (VBK)	Verlagerungswert Cloud-Export für zweites Szenario ¹²³²
$VBK_{Im-Cloud; t}$	Betriebskosten Verlagerung (VBK)	Kosten für das Anlegen neuer Dateien in der Cloud-Ebene über das Internet (je Datei a.) ¹²³³
$VBK_{Ex-Cloud; t}$	Betriebskosten Verlagerung (VBK)	Kosten für den Export der Daten aus der Cloud über das Internet (je Gigabyte im Monat)
$VBK_{Versand-Cloud; t}$	Betriebskosten Verlagerung (VBK)	Kosten für den Import oder Export der Daten der Cloud-Ebene über Versand per Postweg (je Gigabyte und Versandauftrag)

Tab. A5-1: Parameter der Kostenfunktionen

Anhang 5.3 Betriebskosten und Betriebskostenfaktoren im ILM

Die den Parametern der Kostenfunktionen¹²³⁴ zugrunde gelegten Preise¹²³⁵ für die Betriebskosten sowie die jeweiligen Betriebskostenfaktoren werden in diesem Kapitel erörtert. In Tabelle A5-2 sind die Betriebskosten für den Speicherplatz¹²³⁶ der verschiedenen Speicherebenen enthalten.¹²³⁷ Der Preis für ein Gigabyte Cloud-Speicher beträgt im ersten Jahr 2,38 Cent pro Monat.¹²³⁸ Für die Online-Ebene gilt ein Preis von 11,89 Cent pro Monat und Gigabyte¹²³⁹. Hierbei handelt es sich um sogenannten Enterprise-

¹²³¹ Vgl. Kapitel 5.2.3.2

¹²³² Vgl. Kapitel 5.2.3.2

¹²³³ Amazon berechnet die Kosten für das Anlegen neuer Dateien in der Cloud-Ebene je 1.000 Dateien. Vgl. hierzu Kapitel 5.2.3.3.2.1.2.3

¹²³⁴ Vgl. Anhang 5.2

¹²³⁵ Für die Umrechnung der Preise von US-Dollar in Euro wird ein Wechselkurs von 0,7926 (Stand: 01. Oktober 2014, Schlusskurs) zugrunde gelegt. Vgl. hierzu Finanzen.net / Eurokurs / .

¹²³⁶ Vgl. Kapitel 5.2.3.3.1

¹²³⁷ Die in Kapitel 5.2.3.3.2 implementierten Kostenfunktionen berechnen u.a. die Speicherkosten, die im Rahmen der Verwaltung der Daten im ILM entstehen. In der Realität muss neben dem tatsächlich genutzten Speicher zusätzlicher Speicher vorgehalten werden, bspw. für Datensicherungen. Vgl. hierzu Reichman, Whiteley, Chi / File Storage Costs / 4. Die Kosten für zusätzlich vorgehaltenen Speicherplatz werden in den beiden betrachteten Szenarien nicht berücksichtigt.

¹²³⁸ Amazon berechnet für einen Gigabyte Speicher 0,03 US-Dollar Speichergebühren pro Monat (Stand 01. Oktober 2014). Vgl. hierzu Amazon S3 / S3-Preise / . Amazon bietet weiterhin die Möglichkeit, den sogenannten „Reduced Redundancy Storage“ oder den „Glacier Storage“ zu nutzen. Diese Speichermöglichkeiten dienen zur Speicherung von Daten, die für ein Unternehmen geschäftlich wenig relevant sind oder ausschließlich für Archivierungszwecke aufbewahrt werden sollen. Aufgrund der rechtlichen Anforderung an die Ablage und Aufbewahrung der Daten (Vgl. Kapitel 2.4.3.2) werden diese Speichermöglichkeiten in dieser Arbeit nicht betrachtet.

¹²³⁹ Vgl. Null, Lobur / Computer Organization / 765; UC Berkeley / SAN Service / ; Gantz, Reinsel /

Speicher, der eine entsprechende Verfügbarkeit der Daten und einen höheren Datendurchsatz ermöglicht.¹²⁴⁰ Es wird davon ausgegangen, dass der Online-Speicher zur Verwaltung von unstrukturierten Daten bzw. Dokumenten verwendet wird. Wird Online-Speicher bspw. für ERP-Systeme oder die Verwaltung von transaktionsbasierten Datenbanken verwendet, so gelten noch höhere Anforderungen an die Verfügbarkeit und den Durchsatz der Daten. Entsprechend der unternehmensspezifischen Vorgaben liegen die Preise für ein Gigabyte Speicher in diesem Fall bei 2 bis 3 Euro pro Monat.¹²⁴¹ Für die Speicherung der Daten auf der Nearline-Ebene wird angenommen, dass Festplatten mit mittleren Zugriffszeiten zum Einsatz kommen.¹²⁴² Die Preisrelation zwischen den Speichermedien der Online-Ebene und Nearline-Ebene beträgt drei zu eins¹²⁴³ und die Relation zwischen Nearline-Ebene und Offline-Ebene fünf zu eins.¹²⁴⁴ Die Speicherplatzkosten sinken für die jeweilige Speicherebene um 30 Prozent pro Jahr.¹²⁴⁵

Parameter	Kosten [1. Jahr, in €/GB/Monat]	Kosten [2. Jahr, in €/GB/Monat]	Kosten [3. Jahr, in €/GB/Monat]	Kosten [4. Jahr, in €/GB/Monat]	Kosten [5. Jahr, in €/GB/Monat]
SBK Cloud; t	0,0238	0,0166	0,0117	0,0082	0,0057
SBK Online; t	0,1189	0,0832	0,0583	0,0408	0,0285
SBK Nearline; t	0,0396	0,0277	0,0194	0,0136	0,0095
SBK Offline; t	0,0079	0,0055	0,0039	0,0027	0,0019

Tab. A5-2: Betriebskosten Speicherplatz (SBK)

In Tabelle A5-3 sind die Betriebskostenfaktoren enthalten. Der jeweilige Betriebskostenfaktor drückt aus, dass die Kosten für die Administration bzw. Verwaltung der Speichermedien¹²⁴⁶ um ein Vielfaches höher sind, als die Kosten für die Speichermedien

Extracting Value / 4

¹²⁴⁰ Vgl. Sollbach, Thome / ILM Prozessimplementierung / 329-332

¹²⁴¹ Vgl. Minoli / Multicast Applications / 257-258; Mirchandani / Technology Innovations / 63

¹²⁴² Vgl. Kapitel 2.4.2

¹²⁴³ Der Speicher der Online-Ebene kostet das Dreifache als der Speicher der Nearline-Ebene.

¹²⁴⁴ Vgl. SUN / Storage Optimization / 3-4; Turczyk / ILM / 117-118.

¹²⁴⁵ Vgl. hierzu die Annahmen zu den Betriebskosten Kapitel 5.2.3.3.1.2; OECD / Growth and Innovation / 322-323.

¹²⁴⁶ Vgl. hierzu die Kostenarten „Betriebskosten Administration“ und „Allgemeine Betriebskosten“ in

bzw. den Speicherplatz.¹²⁴⁷ Die Kosten für die Administration bzw. Verwaltung der Speichermedien sind unternehmensspezifisch.¹²⁴⁸ Beispielsweise müssen in einem Unternehmen verschiedene Dienstgütevereinbarungen über die Verfügbarkeit und die Antwortzeiten von Informationssystemen eingehalten werden.¹²⁴⁹ Für die Verwaltung der Speichermedien ohne ILM ist ein Betriebskostenfaktor von „vier“ angemessen.¹²⁵⁰ Bezogen auf die Gesamtkosten sind in etwa 20%-30% der Kosten Anschaffungskosten oder Mietkosten für Speicherplatz und 70%-80% sind Kosten für die Administration und Verwaltung der Speichermedien.¹²⁵¹ Bezüglich der Verwaltung der Daten und Speichermedien im ILM ist von einem höheren Betriebskostenfaktor, aufgrund zusätzlicher Betriebskosten,¹²⁵² wie bspw. für die automatisierte Klassifizierung¹²⁵³ der Informationen, für die Miete und den Betrieb von ILM-Software¹²⁵⁴ oder die Verlagerung¹²⁵⁵ der Informationen zwischen verschiedenen Speicherebenen, auszugehen.¹²⁵⁶ Im Rahmen der demonstrierten bzw. durchgeführten Wirtschaftlichkeitsanalyse im ILM wird angenommen, dass der Wert für den Betriebskostenfaktor für Speichermedien ($BKF_{S;t}$) im ersten Jahr „6“ beträgt.¹²⁵⁷ Der Wert für den Betriebskostenfaktor für Cloud-Speicher ($BKF_{C;t}$) beträgt „1,5“. Durch die Nutzung von Cloud-Diensten bzw. der IT-Infrastruktur des Cloud-Anbieters entfallen bspw. die Betriebskosten für Strom oder Wartung der Infrastruktur.¹²⁵⁸ Die Betriebskosten können somit, im Vergleich zur Verwaltung der Daten mit unternehmensinternen Speicherebenen, um 50%-70% gesenkt

Kapitel 5.2.3.3.1, Tabelle 5-1.

¹²⁴⁷ Die Kosten für die Administration der Speichermedien sind um das vier- bis achtfache höher als die Kosten der Datenträger. Vgl. hierzu Glazer / Measuring Information / 99-110; Kaplan, Roy, Srinivasaraghavan / Demand data storage / 1; Forrest, Kaplan, Kindler / Data centers / 1-10; Forrest, Kaplan, Kindler / Energy Efficiency / 1-13; Gantz, Reinsel / Extracting Value / 4; Schadler / Email in the Cloud / 4.

¹²⁴⁸ Vgl. Kaplan, Roy, Srinivasaraghavan / Demand data storage / 1; Forrest, Kaplan, Kindler / Data centers / 1-10; Reichman, Whiteley, Chi / File Storage Costs / 4-11

¹²⁴⁹ Vgl. Hiles / Service Level Agreements / 5-17

¹²⁵⁰ Vgl. Troppens et al. / Storage Networks / 297

¹²⁵¹ Vgl. Anhang 5.3, Tabelle A5-11

¹²⁵² Vgl. Kapitel 5.2.3.3.1

¹²⁵³ Vgl. Kapitel 5.1.3.2

¹²⁵⁴ Vgl. Kapitel 5.2.3.3.1

¹²⁵⁵ Vgl. Kapitel 5.1.2.4

¹²⁵⁶ Vgl. Wittenburg, Trilsbeek, Wittenburg / Archiving and Data Management / 139-140; Reid, Fraser-King, Schwaderer / Managing Data / 2-12

¹²⁵⁷ Dieser Wert repräsentiert den Mittelwert aus den in der Literatur angegebenen Werten 4 bis 8.

¹²⁵⁸ Vgl. Kapitel 3.2.2

werden.¹²⁵⁹ Dennoch entstehen bei der Nutzung von Cloud-Diensten Betriebskosten für bspw. die Ausarbeitung und Einhaltung von Service-Level-Agreements (SLA), die Entwicklung und Anpassung von Schnittstellen zur Nutzung der Cloud-Dienste, die Anpassung der Geschäftsprozesse, die notwendigen Vorbereitungen zum Versand der Speichermedien für den Import und Export der Daten, die Verschlüsselung der Datenübertragung und die Einhaltung rechtlicher Vorschriften.¹²⁶⁰ Der Betriebskostenfaktor für die Verlagerung der Daten ($BKF_{v;t}$) hat einen Wert von „0,5“. Es wird angenommen, dass die Betriebskosten für die Migration bzw. Verlagerung von Daten 50% der Anschaffungs- oder Mietkosten für Speichermedien betragen.¹²⁶¹

Parameter	Betriebskostenfaktor [1. Jahr]	Betriebskostenfaktor [2. Jahr]	Betriebskostenfaktor [3. Jahr]	Betriebskostenfaktor [4. Jahr]	Betriebskostenfaktor [5. Jahr]
$BKF_{S;t}$	6,00	6,30	6,615	6,946	7,293
$BKF_{C;t}$	1,50	1,58	1,654	1,736	1,823
$BKF_{v;t}$	0,50	0,525	0,551	0,579	0,608

Tab. A5-3: Betriebskostenfaktoren (BKF)

Es wird weiterhin angenommen, dass die Betriebskostenfaktoren für Speichermedien, Cloud-Speicher und die Verlagerung der Daten, jährlich um 5 Prozent steigen.¹²⁶² Die Preise für Speicherplatz sinken jährlich. Dennoch steigt der Aufwand für die Administration bzw. Verwaltung der Speichermedien und -ebenen.¹²⁶³ Der Grund hierfür ist die jährlich wachsende Datenmenge von 30%.¹²⁶⁴ Je mehr Daten in einem Unternehmen existieren, desto höher sind die Aufwände, um die Daten zu verwalten.¹²⁶⁵ Weiterhin steigen die Lohn- und Energiekosten.¹²⁶⁶ Die in Tabelle A5-4 dargestellten Administra-

¹²⁵⁹ Vgl. Meinel et al. / Cloud-Speicherdienste / 31-22; Reichman, Whiteley, Chi / File Storage Costs / 4-11

¹²⁶⁰ Vgl. hierzu Baun et al. / Cloud Computing / 39-48; Scharer, Hohmann / Wirtschaftlichkeit IT / 135; Höllwarth / Cloud Migration / 134.

¹²⁶¹ Vgl. Allaire et al. / Costs Data Migrations / 3; Masak / IT-Alignment / 250

¹²⁶² Vgl. Kapitel 5.2.3.3.1.2

¹²⁶³ Vgl. Krcmar / Informationsmanagement / 258

¹²⁶⁴ Vgl. Kapitel 5.2.2.1

¹²⁶⁵ Vgl. Steinhardt / ROI of Storage / 6-7

¹²⁶⁶ Vgl. hierzu Lippold / Unternehmensberatung / 46-47; Elias-Linde / Humanressourcenmanagement / 49-50; Ereik, Schmidt, Schilling / Green-IT / 7; Liedloff, Bromberger / Datenhaltung / 571; Aichele / Smart Energy / 120

tionskosten¹²⁶⁷, bestehend aus den Betriebskosten für die Administration (ABK) und den allgemeinen Betriebskosten (GBK), werden mit Hilfe der Betriebskostenfaktoren berechnet.

Betriebskosten	Kosten [1. Jahr, in €/GB/Monat]	Kosten [2. Jahr, in €/GB/Monat]	Kosten [3. Jahr, in €/GB/Monat]	Kosten [4. Jahr, in €/GB/Monat]	Kosten [5. Jahr, in €/GB/Monat]
Cloud-Ebene	0,0357	0,0375	0,0393	0,0413	0,0434
Online-Ebene	0,7133	0,7490	0,7865	0,8258	0,8671
Nearline-Ebene	0,2378	0,2497	0,2622	0,2753	0,2890
Offline-Ebene	0,0476	0,0499	0,0524	0,0551	0,0578

Tab. A5-4: Betriebskosten Administration (ABK) und allgemeine Betriebskosten (GBK)

Durch Multiplikation des Betriebskostenfaktors für Speicherplatz ($BKF_{S;t}$ und $BKF_{C;t}$) mit den Betriebskosten für Speicherplatz ($SBK_{Cloud;t}$, $SBK_{Online;t}$, $SBK_{Nearline;t}$ und $SBK_{Offline;t}$) werden die Betriebskosten für die Administration (ABK) sowie die allgemeinen Betriebskosten (GBK) berechnet, die durch die Verwaltung der Daten in den verschiedenen Speicherebenen entstehen.¹²⁶⁸ Es gilt die Annahme, dass die Kosten für die Administration und Verwaltung, trotz sinkender Preise für Speicherplatz, steigen.¹²⁶⁹ Hierzu werden die Betriebskosten für die Administration (ABK) und die allgemeinen Betriebskosten (GBK) der vier Speicherebenen (Cloud, Online, Nearline und Offline) für das jeweils erste Jahr berechnet, indem der jeweilige Preis für Speicherplatz (SBK) des ersten Jahres¹²⁷⁰, mit dem jeweiligen Betriebskostenfaktor (BKF) des ersten Jahres¹²⁷¹, multipliziert wird. Der jeweils gültige Preis für den Speicherplatz der vier Speicherebenen im ersten Jahr¹²⁷² dient zusätzlich als Basis- bzw. Referenzwert zur Berechnung der Betriebskosten (ABK und GBK) für die Folgejahre. Dazu wird bspw. im zweiten Jahr der Preis für Speicherplatz des ersten Jahres mit dem im zweiten Jahr gültigen, und gestiegenem, Betriebskostenfaktor, multipliziert. Die Berechnung der Be-

¹²⁶⁷ Vgl. Kapitel 5.2.3.3.1.2

¹²⁶⁸ Vgl. hierzu die Berechnung der Administrationskosten mit Hilfe der Kostenfunktionen Kapitel 5.2.3.3.2.1.1.2 und Kapitel 5.2.3.3.2.1.2.2.

¹²⁶⁹ Vgl. Krcmar / Informationsmanagement / 258

¹²⁷⁰ Vgl. Tabelle A5-2

¹²⁷¹ Vgl. Tabelle A5-3

¹²⁷² Vgl. Tabelle A5-2

triebskosten für die Jahre drei, vier und fünf erfolgt analog dazu. Somit wird im ersten Jahr ein Bezug zwischen Speicherplatzkosten und Betriebskosten hergestellt, da im ersten Jahr die für ILM notwendige Hardware und Software erstmalig beschafft und gemietet wird.¹²⁷³ Basierend auf der nun existierenden Infrastruktur steigen die Betriebskosten jährlich aufgrund wachsender Datenmengen, Lohn- und Energiekosten.¹²⁷⁴

Bezüglich der Betriebskosten für die Administration sind bei der Nutzung der Daten der Cloud-Ebene zusätzlich die Zugriffskosten (ABK_Z) zu berücksichtigen.¹²⁷⁵ Tabelle A5-5 enthält die Kosten für Zugriffe auf die Daten der Cloud-Ebene.¹²⁷⁶

Parameter	Kosten [in €/10.000 Zugriffe, 1.-5. Jahr]
ABK _Z	0,0032

Tab. A5-5: Kosten Zugriffe Cloud

Wenn Daten zwischen den Speichermedien der Speicherebenen Online, Nearline und Offline migriert bzw. verlagert werden, entstehen Betriebskosten für die Verlagerung der Daten (VBK).¹²⁷⁷ Diese sind in Tabelle A5-6 dargestellt.¹²⁷⁸

Betriebskosten	Kosten [1. Jahr, in €/GB/Monat]	Kosten [2. Jahr, in €/GB/Monat]	Kosten [3. Jahr, in €/GB/Monat]	Kosten [4. Jahr, in €/GB/Monat]	Kosten [5. Jahr, in €/GB/Monat]
Online-Ebene	0,0594	0,0624	0,0655	0,0688	0,0723
Nearline-Ebene	0,0198	0,0208	0,0218	0,0229	0,0241
Offline-Ebene	0,0040	0,0042	0,0044	0,0046	0,0048

Tab. A5-6: Betriebskosten Verlagerung Daten (VBK) Speichermedien

¹²⁷³ Vgl. Glazer / Measuring Information / 99-110; Kaplan, Roy, Srinivasaraghavan / Demand data storage / 1; Forrest, Kaplan, Kindler / Data centers / 1-10; Schadler / Email in the Cloud / 4

¹²⁷⁴ Vgl. Tabelle A5-4; Lippold / Unternehmensberatung / 46-47; Elias-Linde / Humanressourcenmanagement / 49-50; Erek, Schmidt, Schilling / Green-IT / 7; Liedloff, Bromberger / Datenhaltung / 571; Aichele / Smart Energy / 120

¹²⁷⁵ Vgl. Kapitel 5.2.3.3.2

¹²⁷⁶ Amazon berechnet 0,004 US-Dollar für die Zugriffe auf die Daten der Cloud-Ebene, je 10.000 Zugriffe (Stand 01. Oktober 2014). Vgl. hierzu Amazon S3 / S3-Preise / .

¹²⁷⁷ Vgl. Kapitel 5.2.3.3.1

¹²⁷⁸ Die Berechnung der Betriebskosten erfolgt analog zur Vorgehensweise zur Berechnung der Betriebskosten für die Administration und den allgemeinen Betriebskosten, unter Verwendung des Betriebskostenfaktors für die Verlagerung der Daten (BKF_{V, t}). Vgl. hierzu Tabelle A5-4.

Werden Daten über das Internet in die Cloud-Ebene übertragen, entstehen ebenfalls Betriebskosten für die Verlagerung der Daten (VBK). Es wird angenommen, dass diese Betriebskosten über die betrachteten fünf Jahre hinweg konstant bleiben. Anbieter von Cloud-Diensten generieren den Großteil ihres Umsatzes mit Übertragungs- bzw. Verlagerungsgebühren, speziell wenn die Daten aus der Cloud-Ebene exportiert werden.¹²⁷⁹ Aus diesem Grund werden die Preise für den Speicherplatz in der Cloud analog zum Preisverfall der Speichermedien pro Jahr vom Cloud-Anbieter gesenkt, die Verlagerungsgebühren bleiben jedoch stabil.¹²⁸⁰ Die Kosten für das Anlegen neuer Dateien in der Cloud-Ebene sind in Tabelle A5-7 dargestellt.¹²⁸¹

Parameter	Kosten [in €/1.000 Dateien, 1.-5. Jahr]
VBK _{Im-Cloud; t}	0,0040

Tab. A5-7: Kosten Anlegen von Dateien

Für den Export der Daten aus der Cloud über das Internet werden die in Tabelle A5-8 abgebildeten Kosten berechnet.¹²⁸²

Parameter	Kosten [in €/Gigabyte, 1.-5. Jahr]
VBK _{Ex-Cloud; t}	0,095

Tab. A5-8: Kosten Datenexport Cloud

Die Gebühren für den Import oder Export der Daten werden vom Cloud-Anbieter auf Basis der Datenmenge¹²⁸³ berechnet, wenn die Speichermedien auf dem Postweg verschickt werden. Die administrativen Aufwände zur bspw. Vorbereitung der Speichermedien für den Postversand und der Anfertigung hierzu notwendiger Backups bzw. Da-

¹²⁷⁹ Vgl. Armbrust et al. / Cloud Computing / 2-25

¹²⁸⁰ Vgl. Amazon / S3 Price Reduction /

¹²⁸¹ Amazon berechnet für das Anlegen neuer Dateien in der Cloud-Ebene 0,005 US-Dollar je 1.000 angelegte Dateien (Stand 01. Oktober 2014). Vgl. hierzu Amazon S3 / S3-Preise / .

¹²⁸² Amazon berechnet für die Übertragung ausgehender Daten über das Internet 0,120 US-Dollar je Gigabyte (Stand 01. Oktober 2014). Vgl. hierzu Amazon S3 / S3-Preise / .

¹²⁸³ Vgl. Anhang 1

tensicherungen sind im Betriebskostenfaktor für Cloud-Speicher berücksichtigt.¹²⁸⁴ Die entsprechenden Betriebskosten sind in den Tabellen A5-9 und A5-10 enthalten.¹²⁸⁵

Parameter	Kosten [in €/Versand, 10 Terabyte]	Kosten [in €/Versand, 13 Terabyte]	Kosten [in €/Versand, 16,9 Terabyte]	Kosten [in €/Versand, 21,97 Terabyte]	Kosten [in €/Versand, 28,56 Terabyte]
VBK _{Versand-Cloud; t}	155,31	175,05	200,70	236,23	279,65

Tab. A5-9: Kosten Versand Speichermedien (10 bis 28,56 TB)

Parameter	Kosten [in €/Versand, 100 Terabyte]	Kosten [in €/Versand, 130 Terabyte]	Kosten [in €/Versand, 169 Terabyte]	Kosten [in €/Versand, 219,70 Terabyte]	Kosten [in €/Versand, 285,61 Terabyte]
VBK _{Versand-Cloud; t}	763,17	1063,10	1354,00	1729,57	2220,21

Tab. A5-10: Kosten Versand Speichermedien (100 bis 285,61TB)

In der Wirtschaftlichkeitsanalyse werden die Kosten bezüglich der Verwaltung der Daten mit und ohne ILM miteinander verglichen.¹²⁸⁶ Zur Berechnung der Betriebskosten für die Administration und der allgemeinen Betriebskosten wird bei der Verwaltung der Speichermedien ohne ILM ein Betriebskostenfaktor¹²⁸⁷ von „vier“ angenommen.¹²⁸⁸ Bezogen auf die Gesamtkosten sind in etwa 20%-30% der Kosten Anschaffungskosten oder Mietkosten für Speicherplatz und 70%-80% sind Kosten für die Administration und Verwaltung der Speichermedien. Es gelten entsprechend die Annahmen aus Tabelle A5-3 zum jährlichen Anstieg des Betriebskostenfaktors.¹²⁸⁹

Parameter	Betriebskosten- faktor [1. Jahr]	Betriebskosten- faktor [2. Jahr]	Betriebskosten- faktor [3. Jahr]	Betriebskosten- faktor [4. Jahr]	Betriebskosten- faktor [5. Jahr]
BKF _{S; t}	4,00	4,20	4,41	4,631	4,862

Tab. A5-11: Betriebskostenfaktor ohne ILM

¹²⁸⁴ Vgl. Tabelle A5-3

¹²⁸⁵ Vgl. Kapitel 5.2.3.3.2.1.2.3; Die Kosten wurden mit Hilfe des „Amazon-Monatsrechners“ berechnet. Der Import der Daten kostet genau so viel wie der Export der Daten. Vgl. hierzu Amazon / Calculator / .

¹²⁸⁶ Vgl. Kapitel 5.3

¹²⁸⁷ Vgl. Tabelle A5-11

¹²⁸⁸ Vgl. Troppens et al. / Storage Networks / 297

¹²⁸⁹ Vgl. Kapitel 5.2.3.3.1.2

Die in Tabelle A5-12 dargestellten Administrationskosten,¹²⁹⁰ bestehend aus den Betriebskosten für die Administration (ABK) und den allgemeinen Betriebskosten (GBK), werden mit Hilfe des Betriebskostenfaktors aus Tabelle A5-11 berechnet.

Betriebskosten	Kosten [1. Jahr, in €/GB/Monat]	Kosten [2. Jahr, in €/GB/Monat]	Kosten [3. Jahr, in €/GB/Monat]	Kosten [4. Jahr, in €/GB/Monat]	Kosten [5. Jahr, in €/GB/Monat]
Online-Ebene	0,476	0,499	0,524	0,551	0,578

Tab. A5-12: Betriebskosten Administration (ABK) und allgemeine Betriebskosten (GBK) ohne ILM

Durch Multiplikation des Betriebskostenfaktors für Speicherplatz ($BKF_{s; t}$) mit den Betriebskosten für Speicherplatz¹²⁹¹ ($SBK_{Online; t}$) werden die Administrationskosten berechnet, die durch die Verwaltung der Daten ohne ILM entstehen. Hierfür werden keine Kostenfunktionen implementiert, da in diesem Fall keine Datenklassifizierung erfolgt und die Daten ausschließlich mit Speichermedien mit kurzen Zugriffszeiten verwaltet werden.¹²⁹²

¹²⁹⁰ Vgl. Kapitel 5.2.3.3.1.2

¹²⁹¹ Es gelten die Annahmen zu den Speicherkosten und dem jährlichen Preisverfall aus Anhang 5.3, Tabelle A5-2.

¹²⁹² Vgl. hierzu die Annahmen zur Berechnung der Administrationskosten Anhang 5.3, Tabelle A5-4.